

Seminararbeit

aus dem Fach Physik

Thema: Wie schlau sind ChatGPT & Co. wirklich?

Verfasser: Alex Eschbaumer
W-Seminar: Wissenschaftsjournalismus
Seminarleiter: Sebastian Bauer
Abgabetermin: 05. November 2024

Erzielte Note: in Worten:
Erzielte Punkte: in Worten:
(einfache Wertung)

Abgabe beim Oberstufenbetreuer am:

Unterschrift des Seminarleiters

Inhaltsverzeichnis

1	Einleitung	1
1.1	Forschungsfrage und Zielsetzung	1
1.2	Methodik und Aufbau	2
2	Die öffentliche Wahrnehmung von Chatbots	3
2.1	Chatbots mit Gefühlen?	3
2.2	Mediale Berichterstattung	4
2.3	Chatbots und IQ-Tests	5
3	Die Funktionsweise von KI-Dialogsystemen	6
3.1	Begriffe und Grundlagen	6
3.2	Transformer und Chatbots	8
3.3	Warum ist das noch nicht AGI?	11
4	Was folgt daraus?	13
4.1	Nur Algorithmen ohne Gefühle	13
4.2	Die falsche Darstellung in den Medien	14
4.3	IQ-Tests mit wenig Aussagekraft	16
5	Fazit: KI kann missverstanden werden	17
	Literaturverzeichnis	18

1 Einleitung

1.1 Forschungsfrage und Zielsetzung

Seit dem 30. November 2022 ist der Chatbot ChatGPT veröffentlicht und frei nutzbar. Nach weniger als zwei Monaten konnte ChatGPT bereits 100 Millionen Nutzer verzeichnen [1, S. 15]. Aufgrund des großen Erfolgs begannen nach kurzer Zeit auch andere Unternehmen wie Google eigene Chatbots zu trainieren. Mittlerweile gibt es immer umfangreichere und leistungsfähigere KI¹ und ihre Entwicklung schreitet rasant voran.

Diese Technologie hat auch einen immer größeren Einfluss auf unsere moderne Gesellschaft. Chatbots gehören für viele Menschen bereits zum Alltag und helfen in der Schule, im Studium und im Beruf, beispielsweise beim Verfassen längerer Texte [2, S. 175, S. 185]. Auf der anderen Seite gibt es aber auch Chatbots, die keinen praktischen Nutzen haben und dazu gedacht sind, sich die Zeit zu vertreiben, indem man z.B. emotionale Gespräche mit ihnen führt. Diese Art von Chatbots spielt vor allem aus sozialer Sicht eine zunehmende Rolle.

Aus diesen Gründen ist es wichtig, diese Technologie richtig einordnen zu können. Allerdings wissen nur wenige wie Chatbots funktionieren, was dazu führt, dass ihre Fähigkeiten oft falsch eingeschätzt werden. Daher soll in dieser Arbeit der Frage nachgegangen werden, welches Bild die Öffentlichkeit von Chatbots hat bzw. was an diesem Bild problematisch ist.

Diese Arbeit verfolgt das Ziel aufzuzeigen, wie es bei der Berichterstattung über KI zu Problemen kommen kann. Außerdem wird die soziale Nutzung von KI-Chatbots diskutiert. Daraus soll erschlossen werden, wie die Öffentlichkeit über KI-Dialogsysteme denkt und warum dadurch ein falscher Eindruck von Chatbots bzw. KI entstehen kann. Aufgrund dessen wird näher auf die Funktionsweise dieser Programme eingegangen. Nach dem Lesen der Arbeit soll man ein grundlegendes Verständnis von Chatbots haben und mit Informationen diesbezüglich aufgeschlossener umgehen können.

¹Künstliche Intelligenz

1.2 Methodik und Aufbau

Im ersten Teil der Arbeit werden verschiedene Beispiele behandelt, die zeigen, dass leicht ein falscher Eindruck von KI, insbesondere Chatbots, entstehen kann. Zunächst wird auf die soziale Bedeutung von KI-Chatbots eingegangen, und die damit verbundenen Konsequenzen und Auswirkungen näher erläutert. Anschließend wird anhand populärwissenschaftlicher Artikel gezeigt, dass die Funktionsweise von KI häufig missverständlich oder ungenau dargestellt wird, was in solchen Quellen leicht zu Missverständnissen führen kann. Darauf folgt eine grundlegende Betrachtung des Themas Chatbots und IQ-Tests. Zunächst wird erläutert, welche Bedeutung IQ-Tests im Zusammenhang mit Chatbots haben. Des Weiteren wird gezeigt, was man bei der Auswertung von Ergebnissen, die Chatbots in IQ-Tests erzielen, beachten muss.

Das darauffolgende Kapitel behandelt die Funktionsweise von Chatbots. Zunächst werden einige wichtige Begriffe geklärt. Es folgt eine Einführung in LLM und GPT. Die grundlegende Funktionsweise von KI-Chatbots wird dabei auf leicht verständliche Weise vermittelt. Mit diesen neuen Erkenntnissen wird anschließend das Thema AGI behandelt. Anhand von ausgewählten Beispielen wird verdeutlicht, warum wir derzeit noch keine Allgemeine Künstliche Intelligenz verwenden. Dieses Thema ist hier besonders passend, da im ersten Teil der Arbeit bereits gezeigt wurde, dass moderne KI häufig falsch dargestellt wird und die Unterscheidung zwischen AGI und ANI (die KI, die wir verwenden) oft schwerfällt.

Im abschließenden Teil der Arbeit sollen Missverständnisse und Problematiken aus dem ersten Teil geklärt werden. Zunächst wird darauf hingewiesen, dass KI derzeit keine menschlichen Gefühle oder Emotionen besitzt. Anschließend werden Missverständnisse aus den zuvor behandelten Artikeln aufgeklärt. Zuletzt wird die Problematik bei der Auswertung von IQ-Tests, die von Chatbots bearbeitet wurden, diskutiert.

2 Die öffentliche Wahrnehmung von Chatbots

2.1 Chatbots mit Gefühlen?

Eine kurze Erklärung vorab: Die meisten kennen bekannte Chatbots wie ChatGPT, die dazu entwickelt wurden, uns bei diversen Aufgaben zu unterstützen. Im Gegensatz dazu steht Character.AI, eines der Unternehmen, die „persönliche“ AEI² anbieten [3]. Das bedeutet, dass die KI darauf ausgelegt ist, soziale Gespräche zu führen. Man kann zwischen verschiedenen Chatbots, die größtenteils fiktive Charaktere repräsentieren, wählen. Sie sind darauf ausgelegt möglichst emotional und menschlich zu reagieren und oft stehen dabei sehr persönliche Themen im Vordergrund, z. B. Einsamkeit, Freundschaft, Empathie oder sogar Liebesbeziehungen.

Das dabei eine emotionale Bindung zu Chatbots entsteht, ist durchaus möglich und dafür reicht ein Chatbot aus, der nur begrenzte soziale Fähigkeiten besitzt [4]. KI ist für viele Menschen also nicht nur eine praktische Anwendung, die im Alltag hilfreich ist, sondern hat für einige Personen auch emotionalen Mehrwert. Ob und warum dies kritisch betrachtet werden muss, wird in 4.1 diskutiert.

Character.AI ist nur eines von mehreren Unternehmen, die persönliche AEI anbieten und das mit Erfolg. Seit September 2022 ist Character.AI frei verfügbar und hat zum jetzigen Zeitpunkt August 2024 bereits über 10 Millionen Downloads bei Google Play. Aufgrund der schnellen Weiterentwicklung von KI und des global stärker werdenden Gefühls von Einsamkeit ist demnach zu erwarten, dass die soziale Bedeutung von Chatbots zunehmen wird [4].

²Artificial Emotional Intelligence

2.2 Mediale Berichterstattung

Die Medien spielen bei der Berichterstattung über KI eine entscheidende Rolle. Sie tragen nicht nur dazu bei, die Öffentlichkeit über aktuelle Entwicklungen und Innovationen zu informieren, sondern sind auch verantwortlich dafür, komplexe technische Zusammenhänge verständlich zu vermitteln. Doch genau hier kann es auch zu Problemen kommen, wodurch eine falsche Auffassung von aktuellen Chatbots entsteht.

Ein Beispiel hierfür ist das Demonstrationsvideo von Google zu ihrem Chatbot Gemini [5]. Das Video zeigt eine Person, die mit dem Chatbot Gemini live interagiert und ihm Fragen stellt. Dabei zeichnet die Person beispielsweise ein Bild von einer Ente und stellt dem Chatbot immer wieder Fragen dazu. Später im Video deutet die Person durch Handbewegungen an, dass sie mit dem Chatbot das Spiel „Schere Stein Papier“ spielen will. Der Chatbot antwortet, dass er die Herausforderung annimmt. Das Video betont dadurch vor allem die Interaktivität zwischen Gemini und der anwesenden Person. Außerdem zeigt das Video einen autonomen Chatbot, der aus einem bestimmten Kontext erschließt welche Antwort gefragt ist.

In vielen Artikeln zu Chatbots wird deren Funktionsweise und Training erklärt. Maria Gramsch erklärt in ihrem Artikel wie ChatGPT funktioniert. Sie fasst sich hierbei kurz und erklärt unter anderem auch das Training von Chatbots. In dem Artikel ist mehrmals zu lesen, dass das Training von ChatGPT stets unvollendet ist und dass der Chatbot aus Konversationen mit den Nutzern immer weiter dazu lernt [6]. Dadurch wird der Eindruck vermittelt, dass ChatGPT wie ein Mensch aus seinen Erfahrungen lernt.

Was genau an den genannten Beispielen zur Berichterstattung falsch ist, wird in 4.2 verdeutlicht. Es steht aktuell eine Vielzahl an Möglichkeiten zur Verfügung, um sich über KI und Chatbots zu informieren. Dadurch kann es schwierig werden eine korrekte Informationsquelle zu wählen, die gut verständlich ist.

2.3 Chatbots und IQ-Tests

Mit der steigenden Zahl an KI Modellen, wird KI-Benchmarking³ immer wichtiger. Außerdem stellt sich die Frage, wie intelligent ein Chatbot verglichen mit einem Menschen ist. IQ-Tests bieten eine Möglichkeit das zu tun, sie sind dazu entworfen Intelligenz so gut wie möglich zu messen und das funktioniert auch bei KI. Jedoch ist es wichtig die Ergebnisse, welche KI in IQ-Tests erzielt, kritisch zu sehen und richtig zu deuten.

Eka Roivainen schreibt in seinem Artikel, wie er mit dem damals neu veröffentlichten ChatGPT-4 einen IQ-Test durchgeführt hat [7]. Er schildert zu Beginn des Artikels, dass er sich dafür interessiere, wie schlau ChatGPT im Vergleich zu menschlichen Standards ist. Im verbalen Teil des verwendeten WAIS-III Tests erklärt er, der Chatbot hat einen Wert von 155 erzielt und ist damit besser als 99,9 % der menschliche Teilnehmer. Der Artikel wurde in anderen Artikeln zitiert, wie beispielsweise von Sandra Blütermann in ihrem Artikel [8]. In einem Interview von Sky News Australia sagt Mo Gawdat, ein ehemaliger Chief Business Officer von Google, ebenfalls, dass ChatGPT einen IQ von 155 habe und damit fast so intelligent wie Albert Einstein sei [9].

Betrachtet man das Ergebnis, was Roivainen mit ChatGPT-4 in ihrem IQ-Test erzielt hat, könnte man meinen, dass KI dem Menschen bereits überlegen ist. Mo Gawdat sagt das sogar schon fast wörtlich in seinem Interview: „Der IQ von ChatGPT-4 wird auf 155 geschätzt, das ist viel schlauer als der durchschnittliche Mensch“ [9]. In 4.3 werden die Ergebnisse, von ChatGPT-4 im WAIS-III Test genauer betrachtet. Darüber hinaus wird die Aussage von Mo Gawdat über die Intelligenz von ChatGPT-4, diskutiert und widerlegt.

³Vergleich von KI-Systemen

3 Die Funktionsweise von KI-Dialogsystemen

3.1 Begriffe und Grundlagen

Um das nächste Kapitel verstehen zu können, müssen zunächst einige wichtige Begriffe und Sachverhalte erklärt werden. Natürlich kann im Rahmen dieser Arbeit keine komplette Erklärung rund um LLM und GPT bis ins Detail gewährleistet werden, darum beziehe ich mich auf das Nötigste, damit man die folgenden Teile der Arbeit gut verstehen kann:

- **Neuronale Netze:** Man kann ein neuronales Netz als eine mathematische Funktion sehen, die eine Anzahl n an Werten annimmt und eine Anzahl m an Werten ausgibt. Diese Werte kann man auch als Eingabe- und Ausgabevektor sehen:

$$f(\vec{v}) = \vec{w}, \quad \text{wobei } \vec{v} \in \mathbb{R}^n \text{ und } \vec{w} \in \mathbb{R}^m$$

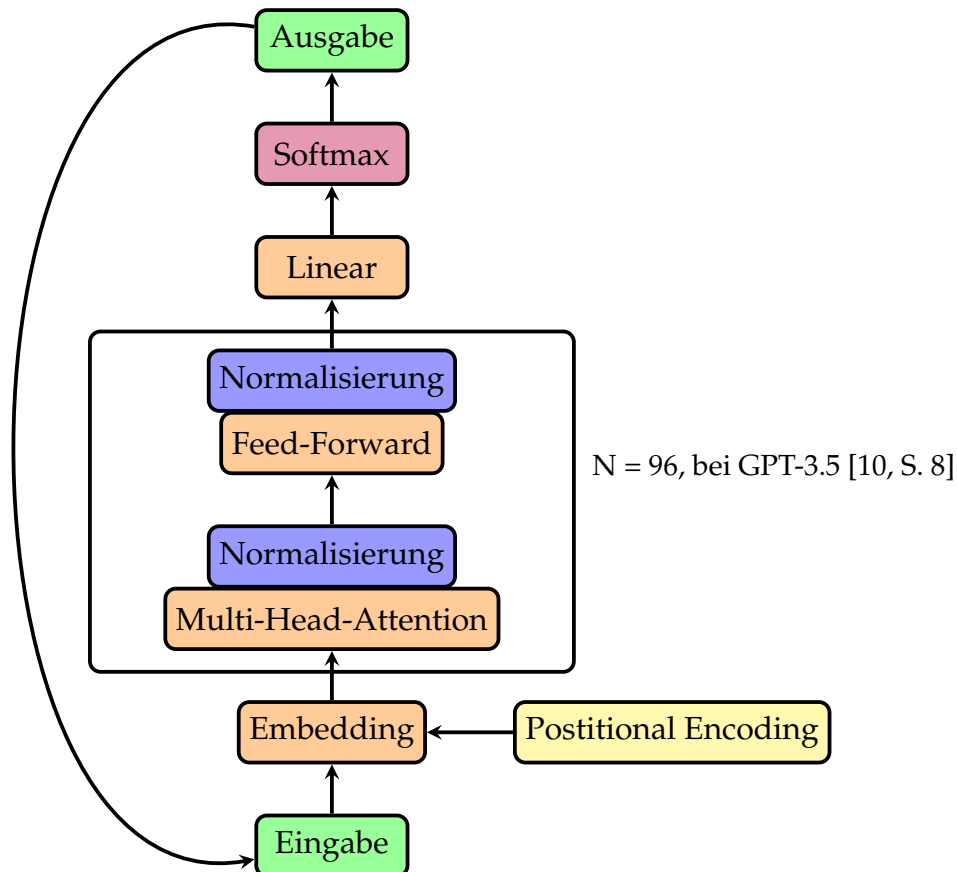
Die Funktion f hat zahlreiche Parameter (175 Milliarden bei GPT-3 [10, S. 8]), die man genau so anpassen will, dass aus unterschiedlichen Eingabevektoren die gewünschten Ausgabevektoren werden und genau hier kommt das Training ins Spiel. Durch Training ist es möglich, die zahlreichen Parameter des Netzwerks so anzupassen, dass sich die Werte der Ausgabevektoren möglichst gut an die gewünschten Werte annähern [11, S. 127f]. Bei der Verwendung eines neuronalen Netzwerks findet keine Veränderung der Parameter statt.

- **Schicht eines Neuronalen Netzwerks:** Ein neuronales Netzwerk ist in mehrere Schichten unterteilt, die verschiedene Bearbeitungsschritte darstellen. Die Eingabevektoren durchlaufen jede Schicht und werden dabei schrittweise in die Ausgabevektoren umgerechnet [11, S. 4].
- **NLP:** Natural Language Processing ist ein Feld der KI, das darauf spezialisiert ist natürliche Sprache zu analysieren, transformieren oder zu generieren.

- **LLM:** Ein Large Language Model ist, wie der Name schon sagt, ein großes KI-Model unserer Sprache. LLMs sind auf NLP spezialisiert und funktionieren unter der Verwendung von neuronalen Netzen.
- **Transformer:** Die Transformer-Architektur ist ein neuronales Netzwerk-Modell, das erstmals im bekannten Forschungsartikel „Attention is all you need“ von Forschern von Google vorgestellt wurde [12]. Transformer in Verbindung mit Chatbots werden in 3.2 genauer erklärt.
- **GPT:** Ein Generative Pre-Trained Transformer ist ein LLM, das sich die Transformer-Architektur zu Nutze macht, um ein deutlich besseres Ergebnis bei der Sprachverarbeitung zu erzielen. Diese GPTs sind der Grund für den Fortschritt im Bereich NLP und sie werden derzeit in fast allen bekannten Chatbot-Modellen, wie ChatGPT und Gemini verwendet [13, S. 2], [14, S. 1].
- **Token:** Wenn ein Chatbot eine Nachricht verarbeitet, dann wird diese zunächst in mehrere kleinere Tokens zerlegt. Tokens helfen bei der Sprachverarbeitung, indem sie Zeichenfolgen darstellen [15]. Zur Vereinfachung kann man sich auch vorstellen, dass ein Token ein Wort repräsentiert. Diesen Tokens wird ein n-dimensionaler Vektor zugewiesen, damit man mit den einzelnen Bestandteilen des Textes Berechnungen durchführen kann.

3.2 Transformer und Chatbots

Im Folgenden wird der schematische Aufbau eines Transformers erklärt. Wie schon in 3.1 erwähnt wurde, ist ein Transformer ein neuronales Netzwerk-Modell. Das heißt, der Transformer nimmt einen Vektor an rationalen Zahlen als Eingabe und gibt ebenso einen Vektor aus.



Die Grafik veranschaulicht den Aufbau eines Transformers, welcher im Folgenden genauer erklärt wird.

Die erste Aufgabe eines Chatbots ist es Text in einzelne Tokens aufzuteilen und dann diesen Tokens einen individuellen Vektor zuzuordnen. Das wird auch als Embedding bezeichnet. Durch das Umwandeln von Tokens in einen Vektor kann das Netzwerk Text mathematisch verarbeiten. Weil die einzelnen Tokens nicht der Reihe, nach sondern unabhängig voneinander und parallel zueinander verarbeitet werden, muss deren Position im Satz mit in den Eingabevektor des Tokens encodiert werden [12, S. 2f]. Das liegt daran, dass bei vielen Wörtern die Position im Satz eine entscheidende Rolle dabei spielt, was der Satz aussagt:

Der Mann jagt den Hund. \neq Der Hund jagt den Mann.

Nach dem Embedding folgt die Multi-Head Attention Schicht. Im Fall eines generativen Transformers wird sich hier der Self-Attention Mechanismus zunutze gemacht [16, S. 2f]. Dabei wird zunächst errechnet wie stark der Zusammenhang zwischen zwei Tokens ist [12, S. 4]. Hat man zum Beispiel die Frage, *Wie nennt man ein großes Haus?* als Eingabe, könnten die errechneten Zusammenhänge für das Wort Haus so aussehen:

Zusammenhang mit Haus:	
<i>Wie</i>	0,046
<i>nennt</i>	0,08
<i>man</i>	0,023
<i>ein</i>	0,12
<i>großes</i>	0,7
<i>Haus</i>	0,031
<i>?</i>	0,001

Weil sich das Adjektiv *großes* auf das Nomen *Haus* bezieht, erkennt ein trainierter Transformer hier einen starken Zusammenhang. Durch diese errechneten Zusammenhänge zwischen den Tokens kann das Netzwerk abwägen, welches Token auf welches andere Token wie viel Einfluss nehmen soll.

Auf die Multi-Head Attention folgt eine Normalisierung, welche dabei hilft die Vektoren zu verarbeiten. Normalisierungen sind zwar Teil der Transformerarchitektur aber nicht entscheidend für ein schematisches Verständnis und werden hier nicht beschrieben[12, S.3].

Der zweite bedeutende Teil eines Transformers nennt sich Feed-Forward-Block. Hier befinden sich c.a. zwei Drittel der Parameter [17]. Parameter bestimmen bekannterweise die Ausgabewerte einer Funktion also eines neuronalen Netzwerks. Im Fall eines Chatbots bestimmen diese Parameter, wie auf einen Satz, den der Transformer als Eingabe bekommt, geantwortet wird. Bei diesen Antworten reicht es aber meistens nicht, einfach nur Zusammenhänge zwischen Wörtern zu erkennen. Das LLM braucht Hintergrundwissen um einen Großteil der Aufgaben zu bearbeiten, die ihm gestellt werden.

Im obigen Satz muss man zum Beispiel nicht nur verstehen, dass es um ein *großes Haus* geht. Ebenfalls muss man wissen, dass alternative Wörter z. B. *Wol-*

kenkratzer oder *Plattenbau* lauten. Dafür braucht man den Feed-Forward-Block. Er kann als Gedächtnis des Transformers gesehen werden. Da sich hier ein Großteil der Parameter befindet, ist hier auch ein maßgeblicher Anteil des Wissens bzw. der Daten, die ein neuronales Netzwerk gelernt hat, gespeichert [17].

Im Attention-Block werden also Zusammenhänge zwischen den Tokens erkannt und dann Informationen zwischen den Tokens übertragen und im Feed-Forward-Block wird dann den eingehenden Tokens bestimmtes Wissen, welches das Netzwerk besitzt, hinzugefügt.

Wichtig ist, dass dieser Ablauf von Attention und Feed-Forward in einem Transformer mehrmals wiederholt wird, damit auch längere komplexe Texte mit komplizierten Zusammenhängen sinnvoll verarbeitet werden können.

Am Ende des Transformers befinden sich noch eine lineare Schicht und eine Softmax Schicht. Diese sind dazu da die Ausgabe des Transformers in einen Vektor umzurechnen, der allen definierten Tokens eine Wahrscheinlichkeit zuweist [11, S. 70]. Darunter wählt man dann das wahrscheinlichste Token aus und fügt es dem Text hinzu.

Die grundlegende Funktionsweise besteht also darin, dass der Transformer die Eingabe in Tokens (welche durch Vektoren repräsentiert wrden) umwandelt, verarbeitet und dann als Ausgabe errechnet, welches Token mit welcher Wahrscheinlichkeit als Nächstes folgt. Natürlich wäre ein Chatbot der nur ein Token als Antwort geben könnte ziemlich nutzlos. Das Problem ist aber einfach zu lösen, indem man dem Transformer dieselbe Eingabe wie zuvor gibt aber noch dazu das neu generierte Token. Das wird auch als Selbstregression bezeichnet und bedeutet, dass der Transformer seine Ausgabe wieder als Eingabe verwendet [12, S. 2]. Diesen Prozess wiederholt man so lange, bis der Transformer ein EOS-Token⁴ generiert und damit den Abschluss seiner Antwort signalisiert [12, S. 2].

⁴End-of-Sequence-Token

3.3 Warum ist das noch nicht AGI?

Bei AGI⁵ handelt es sich um eine Form der KI, die noch nicht existiert. Bis jetzt verwenden wir ausschließlich ANI⁶, das ist KI die für spezifische Aufgaben, wie z. B. Bilderkennung oder Textverarbeitung entwickelt wurde. In 2 wurde herausgearbeitet, dass jedoch oft der Eindruck vermittelt wird, es handle sich bei modernen Chatbots um eine Form von allgemeiner Intelligenz. Im Folgenden wird also herausgearbeitet, weshalb es sich bei den aktuellen Chatbots um ANI handelt. Tatsächlich muss eine KI bestimmte Eigenschaften erfüllen, um als AGI bezeichnet werden zu können [18, S. 8]:

- Eine AGI muss fähig sein ein breites Spektrum an Aufgaben unter verschiedenen Voraussetzungen zu erledigen.
- Eine AGI sollte auch mit Situationen oder Problemen umgehen können, für die sie nicht speziell entwickelt wurde.
- AGI Systeme besitzen die Fähigkeit gelerntes Wissen von einem Problem auf ein anderes zu transferieren.

Zunächst einmal ist ein moderner Chatbot fähig verschiedenste Aufgaben zu erledigen. Man könnte auch argumentieren, dass durch die Bildverarbeitung und Tonverarbeitung, die diese Chatbots oft bereitstellen, verschiedene Voraussetzungen gelten, unter denen ein Chatbot seine Aufgaben erledigt. Die übrigen zwei Aspekte werden jedoch nicht erfüllt.

LLMs sind nicht oder nur begrenzt in der Lage mit Problemen umzugehen, die über das Verarbeiten von Text hinaus gehen. Beispielsweise zeigt Ben Goertzel in seiner Arbeit, dass LLMs ein sehr begrenztes räumliches Vorstellungsvermögen haben [18, S. 32-35]. Wenn Chatbots auf Fragen, die ein solches Vorstellungsvermögen verlangen, richtig antworten wird meistens nur auf gelerntes Wissen zurückgegriffen, d.h. das LLM weiß die Antwort auswendig.

Letztendlich fehlt modernen Chatbots vor allem die Fähigkeit erlerntes Wissen auf andere Probleme zu übertragen. Goertzel schreibt in seiner Arbeit hierzu, dass ein Sprachmodell sein Wissen lediglich als Fallbeispiel abspeichert. Das

⁵Artificial General Intelligence

⁶Artificial Narrow Intelligence

führt dazu das dieses Wissen zwar je nach Bedarf abgerufen werden kann, aber nicht sinnvoll auf andere Situationen transferiert wird. LLMs fehlt eine abstrakte Repräsentation des gelernten Wissens [18, S. 81f].

Insgesamt ist man sich einig, dass es sich bei LLMs wie ChatGPT oder Gemini nicht um AGI handelt. Auch wenn nach außen hin oft der Eindruck vermittelt wird, man spräche mit einem allgemein intelligenten System, ist das nicht der Fall.

4 Was folgt daraus?

4.1 Nur Algorithmen ohne Gefühle

In 2.1 wurde bereits gezeigt, dass KI-Chatbots teils mehr als nur nützliche Assistenten sein können. Jedoch muss die soziale Nutzung von Chatbots kritisch gesehen werden. Dafür gibt es folgende Gründe:

Zunächst ist es wichtig zu verstehen, dass aktuelle Chatbots keine Gefühle empfinden, auch wenn es nach außen hin so wirkt. Wie bereits in 3.2 gezeigt wurde handelt es sich bei modernen Chatbots um LLMs, also Programme, die durch Berechnung Texte generieren. Das LLM wird zunächst anhand von Beispielttext trainiert. Will man also, dass ein Chatbot fähig ist ein soziales Gespräch zu führen würde man ihn beispielsweise mit passenden Chatverläufen trainieren. Dadurch lernt der Chatbot zwar während sozialen Konversationen eine passende Antwort zu geben, jedoch versteht er dessen Bedeutung nicht [18, S. 35-38]. Das heißt also, dass Chatbots lediglich anhand von Trainingsdaten lernen so zu tun als hätten sie Gefühle.

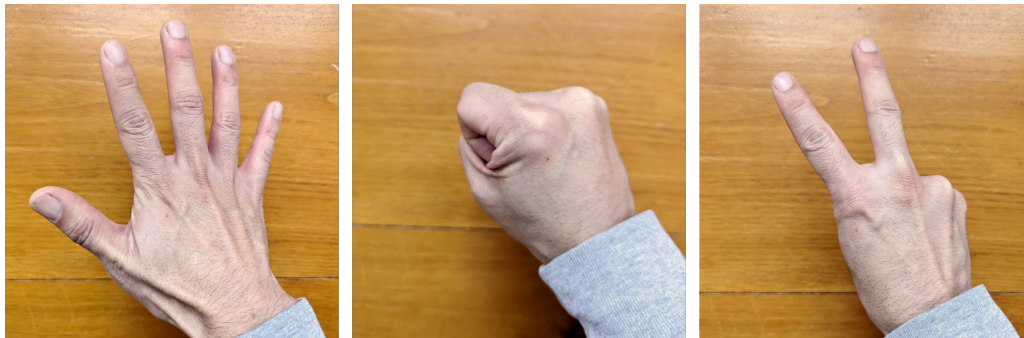
Abgesehen von der Tatsache, dass Chatbots nicht in der Lage sind Gefühle zu empfinden, kann die vermehrte Nutzung von KI anstelle des Menschen zu einer Dehumanisierung führen. Indem man einem Chatbot menschliche Eigenschaften zuschreibt und Programme auf dasselbe Level wie reale Menschen stellt, verliert die soziale Interaktion mit echten Menschen an Bedeutung. Außerdem wird die soziale Komponente des Menschen dadurch ersetzbar gemacht, wodurch sie an Wert verliert [4].

Natürlich gibt es auch positive Aspekte, die eine Nutzung von KI in sozialen Bereichen mit sich bringt. Die aufgeführten Gründe zeigen aber, dass man zwischen der Verwendung von AEI und Interaktionen mit realen Menschen differenzieren sollte.

4.2 Falsche Darstellung in den Medien

Die genannten Beispiele aus 2.3 sollen im folgenden Teil erneut aufgegriffen und richtig gestellt werden. Darüber hinaus werden die technischen Hintergründe zu den Beispielen genauer beleuchtet.

Zu Beginn wurde auf die Gemini Demo eingegangen. Im Video zur Demo wurde vor allem betont, wie Interaktiv der Chatbot mit der anwesenden Person zusammenarbeitet und es entsteht der Eindruck, dass der Chatbot in Echtzeit auf die Person reagiert. In Wahrheit hat der Chatbot aber nur Bilder und Text mit zusätzlichen Hinweisen als Eingabe bekommen [19]. Beispielsweise wurde dem Chatbot das folgende Bild gezeigt, mit dem Hinweis, dass es sich um ein Spiel handelt:



Der Chatbot hat also nicht durch eine Echtzeitaufnahme anhand von Handbewegungen erkannt, dass es sich um das Spiel „Schere Stein Papier“ handelt, sondern durch ein Bild mit zusätzlichem Hinweis. Dieses Schema zieht sich durch das gesamte Video und es wird nicht erwähnt, dass der Chatbot ausschließlich Bilder mit zusätzlichem Text als Eingabe erhält. Obwohl in der Beschreibung des Videos ein Artikel verlinkt ist, indem die Erstellung verdeutlicht wird, hat das Video bei vielen Leuten für falsche Erwartungen und letztendlich für Kritik gesorgt.

Das zweite Beispiel bezog sich auf einen Artikel, indem die Funktionsweise und das Training von Chatbots erklärt wird. Im Artikel wird mehrmals erwähnt, dass Chatbots aus den Konversationen mit den Nutzern lernen [6]. In 3.2 wurde genauer erklärt, wie ein Chatbot auf Nutzeranfragen antwortet. Während der Chatbot eine Antwort generiert findet kein Training statt, da die Parameter des Transformers nicht verändert werden. Die Aussage, dass der Chatbot aus den Fragen der Nutzer lernt ist also falsch.

Beide genannten Fälle sorgen dafür, dass KI besser dargestellt wird, als sie eigentlich ist. Dafür gibt es mehrere Hintergründe, wie z. B. Marketing im Fall von Gemini oder mangelndes Wissen. Problematisch ist, dass vor allem Personen, die nur wenig Verständnis von KI besitzen, durch solche Informationsquellen leicht einen falschen Eindruck davon bekommen können, wozu Chatbots fähig sind.

4.3 IQ-Tests mit wenig Aussagekraft

Kapitel 2.3 hat IQ-Tests im Zusammenhang mit Chatbots behandelt. Dabei wurde erklärt, wie ChatGPT-4 in einem IQ-Test eine Punktezahl von 155 erzielte. Im Folgenden soll diskutiert werden, warum dieses Ergebnis nicht mit dem eines Menschen vergleichbar ist und kritisch gesehen werden muss.

Zum einen wurde nur ein verbaler IQ-Test durchgeführt. Natürlich hat man bei einem textbasierten Chatbot keine andere Möglichkeit Fragen zu stellen. Roivainen schreibt in seinem Artikel, dass der verbale und allgemeine IQ stark zusammenhängen. Er erklärt, dass ChatGPT nach menschlichen Maßstäben sehr intelligent ist [7]. Der Zusammenhang zwischen verbalen und allgemeinen IQ wurde jedoch bei Menschen und nicht bei Chatbots beobachtet. Darum ist diese Folgerung so nicht möglich.

Zudem sind die Fragen, die der Chatbot beantwortet hat, nicht geeignet, wenn es darum geht festzustellen, wie intelligent eine KI ist. Bekommt ein Mensch z. B. die Frage „Was haben Harry Potter und Bugs Bunny gemeinsam“ [7], dann muss er nachdenken, wodurch man einen sinnvollen Zusammenhang zwischen Bugs Bunny und Harry Potter herstellt. Bekommt hingegen ein Chatbot diese Frage, dann kann er sie meistens anhand der Trainingsdaten beantworten. Ein Chatbot kommt nicht auf dieselbe Weise zu einer Antwort wie ein Mensch. Das wird in 3.2 deutlich. Daraus folgt auch, dass man die Ergebnisse von Menschen und Chatbots in solchen IQ-Test nicht sinnvoll vergleichen kann. Somit ist die Folgerung von Roivainen, dass ChatGPT nach menschlichen Maßstäben sehr intelligent ist, nicht zutreffend [7]. Außerdem wurde in 3.3 gezeigt, dass es sich bei Chatbots nicht um allgemeine Intelligenz handelt. Ihnen fehlen essenzielle Eigenschaften, die eine zentrale Rolle bei der Definition von allgemeiner Intelligenz spielen.

Das Ergebnis, welches ChatGPT-4 im WAIS-III IQ-Test erzielt hat, wurde mehrmals zitiert. Wie im Artikel selbst wurde die Vergleichbarkeit zwischen Mensch und Chatbot einfach vorausgesetzt [9] oder nicht kritisch hinterfragt [8]. Durch das aufgeführte Beispiel wird deutlich, wie leicht Ergebnisse, die KI in IQ-Tests erzielt, missverstanden werden können. Allgemein ist es sehr schwer die Intelligenz eines Chatbots mit der eines Menschen zu vergleichen.

5 Fazit: KI kann missverstanden werden

Zusammenfassend lässt sich feststellen, dass es durchaus möglich ist, eine falsche Auffassung von KI zu bekommen. Dazu braucht es nicht einmal direkt falsche Informationen, sondern vor allem eine mangelnde Intuition, was tatsächlich hinter den aktuellen Chatbots steckt. Es ist wichtig grundlegend zu verstehen wie Chatbots funktionieren, um Neuigkeiten und Informationen über Chatbots richtig einordnen zu können. Das wurde in der Arbeit durch folgende drei Beispiele verdeutlicht.

Chatbots stellen für einige Menschen einen sozialen Mehrwert dar. Personen, die Chatbots für diese Zwecke verwenden sollten sich bewusst sein, dass Chatbots keine Empfindungsfähigkeit besitzen. Zudem sollte man klar zwischen der sozialen Interaktion mit Chatbots und der mit Menschen unterscheiden.

Die Medien spielen eine zentrale Rolle dabei, die Öffentlichkeit über Neuigkeiten zum Thema KI und Chatbots zu informieren. Das heißt jedoch nicht, dass Informationen, die in den Medien vermittelt werden, richtig sind. Mit einem grundlegenden Verständnis ist es aber deutlich leichter zwischen richtigen und falschen Informationen zu unterscheiden.

Auch wenn ein Chatbot nach außen hin sehr intelligent wirkt und man das auch mit einem IQ-Test messen kann, muss man die Ergebnisse von Chatbots und Menschen unterscheiden. Chatbots haben zahlreiche Schwächen und Einschränkungen, die sich aber nicht direkt bemerkbar machen.

Natürlich konnte im Rahmen dieser Arbeit nur ein Teil der Faktoren, die die öffentliche Wahrnehmung von Chatbots beeinflussen, behandelt werden. Es gibt auch zahlreiche gute Quellen, um sich über KI zu informieren. Die Arbeit hat aber deutlich gemacht, dass man leicht auf falsche Informationen stoßen kann.

Letztendlich werden Chatbots bzw. KI immer mehr in unser modernes Leben integriert. Aus diesem Grund wird auch ein Verständnis dieser Technologie immer wichtiger um damit verantwortungsvoll umgehen zu können und Missverständnisse zu vermeiden. Darum bietet diese Arbeit eine vereinfachte Erklärung von GPTs, die den Leserinnen und Lesern einen Einblick in das Thema Künstliche Intelligenz gibt und ihnen dabei hilft, die Funktionsweise und die Grenzen dieser Technologie besser zu verstehen.

Literaturverzeichnis

- [1] Manfred Spitzer. *Künstliche Intelligenz, Dem Menschen Überlegen - wie KI uns rettet und bedroht*. Bd. 1. München: Droemer Verlag, Nov. 2023. ISBN: "978-3-426-44850-2".
- [2] Dr. A. Shaji George, A. S. Hovan George und A. S. Gabrio Martin. „ChatGPT and the Future of Work: A Comprehensive Analysis of AI’s Impact on Jobs and Employment“. In: *Partners Universal International Innovation Journal* 1.3 (Juni 2023), S. 154–186. DOI: 10.5281/zenodo.8076921. URL: <https://www.puiij.com/index.php/research/article/view/57>.
- [3] Noam Shazeer und Daniel De Freitas. *Character.AI*. Auch als mobile App verfügbar. Aug. 2024. URL: <https://character.ai/>.
- [4] Angelina Chen u. a. „Feels Like Empathy: How “Emotional” AI Challenges Human Essence“. In: *Australasian Conference on Information Systems*. 2024.
- [5] Google. *The capabilities of multimodal AI | Gemini Demo*. Youtube. 2023. URL: <https://www.youtube.com/watch?v=UIZAiXYceBI>.
- [6] Maria Gramsch. „Künstliche Intelligenz: Wie funktioniert eigentlich ChatGPT?“. In: *BASIC thinking* (Mai 2023).
- [7] Eka Roivainen. „I Gave ChatGPT an IQ Test. Here’s What I Discovered“. In: *Scientific American* (März 2023). URL: <https://www.scientificamerican.com/article/i-gave-chatgpt-an-iq-test-heres-what-i-discovered/>.
- [8] Sandra Blütermann. „ChatGPT besteht IQ-Test“. In: *Computer Bild* (März 2023). URL: <https://www.computerbild.de/artikel/cb-News-Internet-ChatGPT-besteht-IQ-Test-35523461.html>.
- [9] Sky News Australia. *Chat-GPT ‘much smarter’ than the average human: Mo Gawdat*. Youtube. 2023. URL: <https://www.youtube.com/watch?v=BKf7qrcN0X8>.
- [10] OpenAI. „Language Models are Few-Shot Learners“. Version 4. In: (Juni 2020). arXiv Identifier: 2005.14165v4. URL: <https://arxiv.org/abs/2005.14165v4>.

- [11] Michael A Nielsen. *Neural networks and deep learning*. Bd. 25. Determination press San Francisco, CA, USA, 2015. URL: <https://ise.ncsu.edu/wp-content/uploads/sites/9/2022/08/neuralnetworksanddeeplearning.pdf>.
- [12] Ashish Vaswani. „Attention is all you need“. In: *Advances in Neural Information Processing Systems* (2017). URL: <https://user.phil.hhu.de/~cwurm/wp-content/uploads/2020/01/7181-attention-is-all-you-need.pdf>.
- [13] Google Gemini Team. *Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context*. Techn. Ber. Version 4. arXiv Identifier: 2403.05530v4. Aug. 2024. URL: <https://arxiv.org/abs/2403.05530v4>.
- [14] OpenAI. *GPT-4 Technical Report*. Techn. Ber. Version 6. arXiv Identifier: 2303.08774v6. März 2024. URL: <https://arxiv.org/abs/2303.08774v6>.
- [15] Sabrina J Mielke u. a. „Between words and characters: A brief history of open-vocabulary modeling and tokenization in NLP“. Version 1. In: (2021). arXiv Identifier: 2112.10508. URL: <https://arxiv.org/abs/2112.10508>.
- [16] Richard E. Turner. „An Introduction to Transformers“. Version 5. In: (Feb. 2024). arXiv Identifier: 2304.10557v5. URL: <https://arxiv.org/abs/2304.10557v5>.
- [17] Mor Geva. „Transformer Feed-Forward Layers Are Key-Value Memories“. Version 2. In: (Sep. 2021). arXiv Identifier: 2012.14913v2. URL: <https://arxiv.org/abs/2012.14913v2>.
- [18] Ben Goertzel. „Generative AI vs. AGI: The Cognitive Strengths and Weaknesses of Modern LLMs“. Version 1. In: (2023). arXiv Identifier: 2309.10371. URL: <https://arxiv.org/abs/2309.10371>.
- [19] Alexander Chen. „How it’s Made: Interacting with Gemini through multi-modal prompting“. In: (Dez. 2023). Aus Google for Developers. URL: <https://developers.googleblog.com/en/how-its-made-interacting-with-gemini-through-multimodal-prompting/>.

Ich erkläre hiermit, dass ich diese Arbeit ohne fremde Hilfe angefertigt habe und nur die im Literaturverzeichnis angegebenen Quellen und Hilfsmittel benutzt habe.

_____, den _____