

Seminararbeit

aus dem Fach Physik

Thema: Wie schlau sind ChatGPT & Co. wirklich?

Verfasser: Alex Eschbaumer
W-Seminar: Wissenschaftsjournalismus
Seminarleiter: Sebastian Bauer
Abgabetermin: 07. 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	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 2023 ist der Chatbot ChatGPT veröffentlicht und frei nutzbar. Nur nach weniger als zwei Monaten hatte ChatGPT bereits 100 Millionen Nutzer [1, S. 15]. Wegen des großen Erfolges begannen schon nach kurzer Zeit weitere Unternehmen wie Google ihre eigenen Chatbots zu entwickeln. Mittlerweile gibt es immer umfangreichere sowie leistungsfähigere KI und deren Entwicklung schreitet unglaublich schnell voran.

Diese Technologie hat einen immer größer werdenden Einfluss auf unsere moderne Gesellschaft. Chatbots sind für viele Menschen bereits Teil ihres Alltags und helfen in der Schule oder im Studium, aber auch im Beruf, beispielsweise dann, wenn es darum geht längere Texte zu verfassen [2, S. 175, S. 185]. Auf der andere Seite gibt es sogar Chatbots die keinen praktischen Nutzen haben und dazu gedacht sind, dass man sich mit ihnen die Zeit vertreibt, indem man z. B. emotionale Gespräche mit ihnen führt. Diese Art der Chatbots sollen als eine Art KI-Freund dienen, sie sind vor allem in Form von Anwendungen aus dem Play Store & App Store bekannt.

Es ist wichtig so eine Technologie richtig einordnen zu können. Jedoch wissen die wenigsten wie Chatbots funktionieren, was dazu führt, dass deren Fähigkeiten oft falsch eingeschätzt werden. Darum soll in dieser Arbeit der Frage nachgegangen werden, welches Bild die Öffentlichkeit von Chatbots hat, bzw. was an diesem Bild problematisch sein kann.

Diese Arbeit verfolgt das Ziel aufzuzeigen, welcher Eindruck von Chatbots in den Medien vermittelt wird. Daraus sollen Schlüsse gezogen werden, wie Öffentlichkeit über KI denkt. Außerdem soll die Arbeit verständlich erklären, warum der Eindruck von Chatbots bzw. KI oft fehlerhaft ist. Dazu 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.

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, vermittelt wird. Zunächst wird auf die soziale Bedeutung bestimmter KI-Chatbots eingegangen, und die damit verbundenen Konsequenzen und Folgen werden genauer erläutert. Anschließend wird anhand populärwissenschaftlicher Artikel aufgezeigt, dass die Funktionsweise von KI häufig missverständlich oder ungenau dargestellt wird, was in solchen Quellen leicht zu Missverständnissen führen kann. Daraufhin erfolgt 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 erzielen, beachten muss.

Der zweite Teil der Arbeit beschäftigt sich mit der Funktionsweise von Chatbots. Zunächst werden einige wichtige Begriffe geklärt, gefolgt von einer Einführung in das LLM und den GPT. Die grundlegende Funktionsweise von KI-Chatbots wird dabei leicht verständlich vermittelt. Mit diesem neuen Wissen wird anschließend das Thema AGI¹ behandelt. Anhand ausgewählter Beispiele wird verdeutlicht, warum zum gegenwärtigen Zeitpunkt noch keine AGI entwickelt wurde. Das Thema AGI ist hier besonders relevant, da im ersten Teil der Arbeit bereits gezeigt wurde, dass moderne KI häufig falsch dargestellt wird und die Unterscheidung zwischen AGI und ANI² oft schwerfällt.

Im abschließenden Teil der Arbeit sollen Missverständnisse und Problematiken aus dem ersten Teil geklärt werden. Zunächst wird darauf hingewiesen, dass KI derzeit weder intelligent ist noch menschliche Gefühle oder Emotionen besitzt. Anschließend werden Missverständnisse aus den zuvor behandelten Artikeln aufgeklärt. Abschließend wird die Problematik von IQ-Tests, die von Chatbots bearbeitet wurden, diskutiert.

¹Artificial General Intelligence

²Artificial Narrow Intelligence

2 Die öffentliche Wahrnehmung von Chatbots

2.1 Chatbots mit Gefühlen?

Eine kurze Erklärung vorab: Die meisten kennen vermutlich 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“ AEI³ anbieten [3]. Das bedeutet, dass die KI dazu entworfen wurde, soziale Gespräche zu führen. Man kann zwischen verschiedenen Chatbots, die meist 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 Verbindung zu Chatbots entstehen kann ist durchaus möglich und dafür reicht ein Chatbot aus, der nur begrenzte soziale Fähigkeiten hat [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 wieso das kritisch gesehen werden muss wird 4.3 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. Auf Grund 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].

³Artificial 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 damit vor allem die Interaktivität zwischen Gemini und der anwesenden Person. Außerdem zeigt das Video einen autonomen Chatbot, der aus einem bestimmten Context Schlüsse zieht, 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 zwei mal 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 passende Informationsquelle zu wählen, die gut verständlich ist.

2.3 Chatbots und IQ-Tests

Mit der steigenden Zahl an KI Modellen, wird KI-Benchmarking⁴ 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 ihrem Artikel, wie er mit dem damals neu veröffentlichten ChatGPT-4 einen IQ-Test durchgeführt hat [7]. Sie schildert zu Beginn des Artikels, dass sie sich dafür interessiere, wie schlau ChatGPT im Vergleich zu menschlichen Standards ist. Im verbalen Teil des verwendeten WAIS-III Tests erklärt sie, der Chatbot hat einen Wert von 155 erzielt und ist damit besser als 99,9% der menschliche Teilnehmer. Der Artikel wurde mehreren anderen Artikeln zitiert, wie beispielsweise von Sandra Blütermann in ihrem Artikel [8]. In einem Interview von Sky News Australia erklärt 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 ?? werden die Ergebnisse, von ChatGPT-4 im WAIS-III Test genau betrachtet. Darüber hinaus wird die Aussage von Mo Gawdat über die Intelligenz von ChatGPT-4, diskutiert und widerlegt.

⁴Vergleich 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 nicht eine komplette Erklärung rund um LLM und GPT bis ins Detail gewährleistet werden, darum beziehe ich mir hier auf das Nötigste, um die folgenden Teile der Arbeit gut verstehen zu können:

- **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 einzustellen, dass sich die Werte der Ausgabevektoren möglichst gut an die gewünschten Werte annähern [11, S. 127f]. Bei der Verwendung eines neuronalen Netzwerks findet kein Training statt.

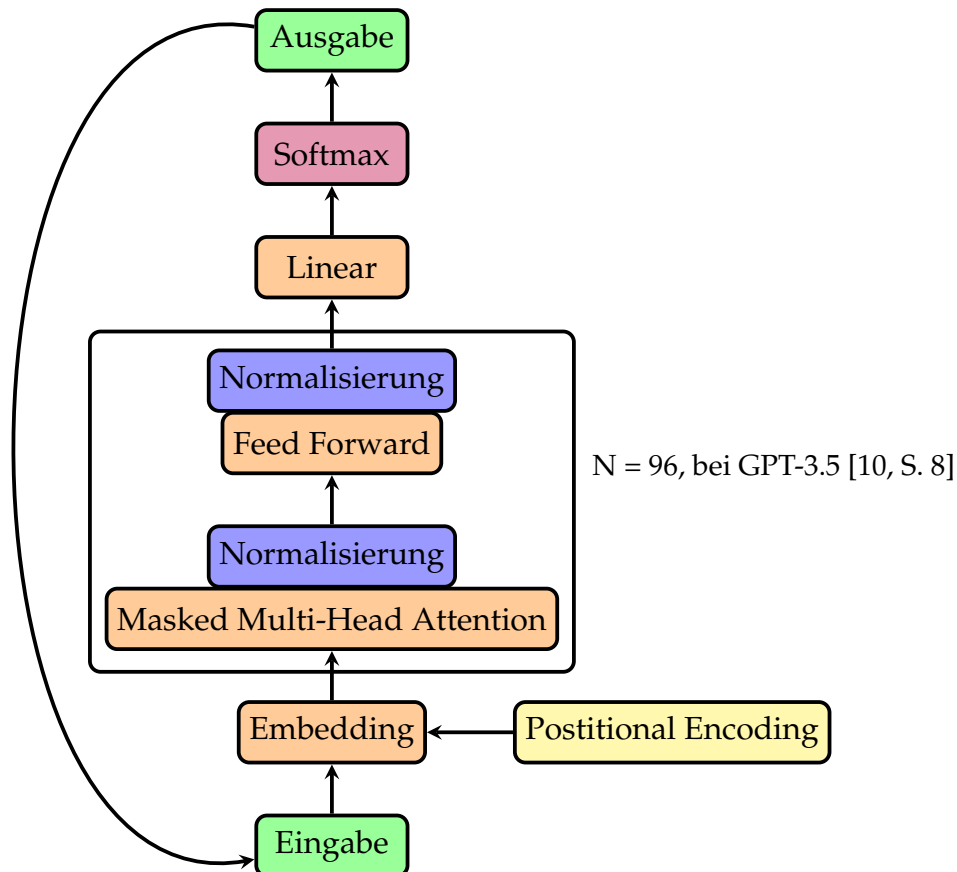
- **Schicht eines Neuronalen Netzwerks:** In einem neuronalen Netzwerk gibt es mehrere Schichten, die einzelne Bearbeitungsschritte darstellen. Die Eingabevektoren durchlaufen jede Schicht und werden dabei in die Ausgabevektoren umgerechnet [11, S. 4].
- **NLP:** Natural Language Processing ist ein Feld der KI, dass 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 Transformerarchitektur ist ein neuronales Netzwerk-Modell, das erstmals im bekannten Forschungsartikel „Attention is all You Need“ von Forschern bei Google vorgestellt wurde [12]. Transformer in Verbindung mit Chatbots werden in 3.1 genauer erklärt.
- **GPT:** Ein Generative Pre-Trained Transformer ist ein LLM das sich die Transformerarchitektur 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 in Wörtern darstellen [15]. Zur vereinfachung kann man sich auch vorstellen, dass ein Token ein Wort räsentiert.

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 erste Aufgabe eines Chatbots ist es also Text zunächst 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 von einander und parallel zueinander verarbeitet werden, muss deren Position im Satz mit in den Eingabevektor encodiert werden [12, S. 2f]. Das liegt daran, dass bei vielen Wörtern die Position im Satz eine entscheidende Rolle dafür spielt, was der Satz aussagt:

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

Die grundlegende Funktionsweise besteht darin, dass der Transformer diese Tokens in Zahlen 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 die Tokens der vorherigen Eingabe und dazu die neu generierten Tokens gibt. Das wird auch als Selbstregression bezeichnet und bedeutet, dass der Transformer seine Ausgabe wieder als Eingabe verwendet [12, S. 2].

Nach dem Embedding folgt die eine Masked Multi-Head Attention Schicht⁵. 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 zu 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.

Nach der 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 genauer beschrieben[12, S.3].

⁵Teil eines neuronalen Netzwerks

Der zweite bedeutende Teil eines Transformers folgt direkt auf den attention Block und nennt sich Feed Forward Block. Im Feed Forward Block befinden sich c.a. zwei drittel der Parameter [17]. Parameter bestimmen bekanntlicher weise die Ausgabewerte einer Funktion oder 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 meistens aber nicht aus, einfach nur Zusammenhänge zwischen Wörtern zu erkennen, wie oben erklärt, sondern man benötigt Hintergrundwissen zum Thema für eine konstruktive Antwort.

Im obigen Satz muss man zum Beispiel nicht nur verstehen, das es um ein *großes Haus* geht sondern auch wissen welche alternativen Wörter es dafür gibt, um die Frage zu beantworten. Dafür braucht man den Feed Forward Block, er kann als Gedächtnis des Transformers gesehen werden. Da sich hier ein großer Teil der Parameter befindet ist hier auch ein großer Teil 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 noch auf die eingehenden Tokens mit bestimmten Wissen, welches das Netzwerk besitzt reagiert.

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 Transformer befinden sich noch eine Lineare und eine Softmax Schicht. Sie sind dazu da die Ausgabe des Transformers in einen Vektor zu überführen, der allen definierten Tokens eine Wahrscheinlichkeit zuweist [11, S. 70]. Darunter wählt man dann meist das Token mit der höchsten Wahrscheinlichkeit aus und fügt es dem Text hinzu. Diesen Prozess wiederholt man dann so lange, bis der Transformer ein EOS-Token⁶ generiert und damit den Abschluss seiner Antwort signalisiert [12, S. 2].

⁶End-of-Sequence-Token

3.3 Warum ist das noch nicht AGI?

Bei AGI handelt es sich um eine Form von KI, die noch nicht existiert. Bis jetzt verwenden wir ausschließlich ANI, das ist KI die für spezifische Aufgabe, wie z. B. Bilderkennung entwickelt wurde. In 2 wurde herausgearbeitet, dass jedoch oft der Eindruck vermittelt wird, es handele sich bei modernen Chatbots um eine Form von AGI. Im Folgenden wird also herausgearbeitet wieso es sich bei den aktuellen Chatbots um ANI handelt. Allgemein 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 Aufgabe 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 Aufgabe 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 um zu gehen, 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 fragen, die ein solches Vorstellungsvermögen verlangen, richtig beantwortet wird meistens nur auf gelerntes Wissen zurück gegriffen, d.h. das LLM weis die Antwort auswendig.

Letzendlich 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 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 einer allgemeinen 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 in der Gesellschaft teils mehr als nur nützliche Assistenten sind. 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 Beispieltext 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.

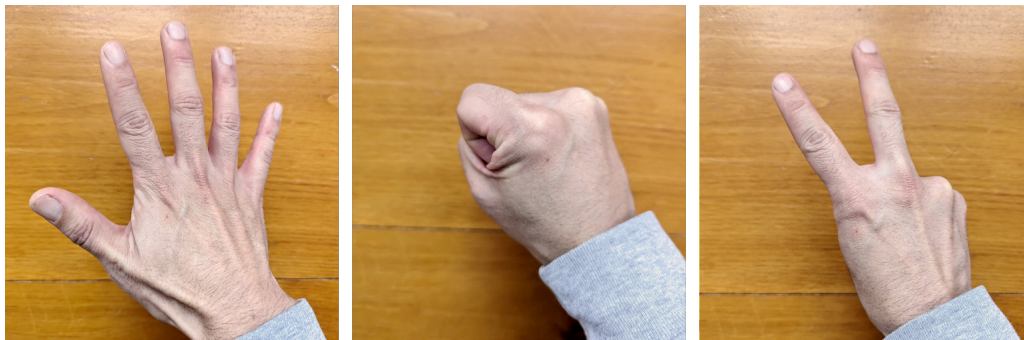
Abgesehen von der Tatsache, dass Chatbots nicht in der Lage sind Gefühle zu empfinden, kann die vermehrte Nutzung von KI anstelle des Menschen zu einer Entmenschlichung führen. Indem man einem Chatbot menschliche Eigenschaften zuschreibt und Programme auf das gleiche Level wie reale Menschen stellt, 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 echtzeit Aufnahme anhand von Handbewegungen erkannt, dass es sich um das Spiel „Schere Stein Papier“ handelt, sondern durch ein Bild mit zusätzlichen Hinweis. Dieses Schema zieht sich durch das gesamte Video und es wird nicht erwähnt, dass der Chatbot nur Bilder mit zusätzlichem Text als Eingabe bekommt. Obwohl in der Beschreibung des Videos ein Artikel verlinkt ist, indem die Erstellung verdeutlicht wird, hat das Video bei vielen Leuten für flasche 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, wie Chatbots aus den Konversationen mit den Nutzern lernen [6]. In 3.2 wurde genauer erklärt, wie ein Chatbot auf Nutzeranfrage antwortet. Während der Chatbot eine Antwort generiert findet kein Training statt, da die Parameter des Transformers nicht verändert werden. Die Aussage, dass das der Chatbot aus den Fragen der Nutzer lernt ist also flasch.

Beide genannten Fälle sorgen dafür, dass KI besser dargestellt wird, als sie ei-

gentlich ist. Dafür gibt es mehrere Hintergründe, wie z. B. Marketing im fall von Gemini oder mangelndes Wissen. Problematisch ist, dass vorallem Personen, die nur wenig Verständnis von KI besitzen, durch solche Informationsquellen leicht einen flaschen Eindruck davon bekommen können, zu was 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 ihrem Artikel, dass ein verbale IQ und der allgemeine IQ stark zusammenhängen. Sie schreibt, 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 fest zu stellen, 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 die selbe Weise zu einer Antwort wie ein Mensch, dass 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 AGI handelt. Ihnen fehlen essentielle 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 wirklich 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 auch einen sozialen Mehrwert dar. Personen, die Chatbots für diese Zwecke verwenden sollten sich bewusst sein, dass Chatbots nicht empfindungsfähig sind. 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 falsche und richtige 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.

Letzendlich werden Chatbots und KI immer mehr in unser modernes Leben integriert. Darum wird auch ein Verständnis dieser Technologie immer Wichtiger um damit verantwortungsvoll umgehen zu können und Missverständnisse zu vermeiden.

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).
- [8] Sandra Blütermann. „ChatGPT besteht IQ-Test“. In: *Computer Bild* (März 2023).
- [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 das Exposé ohne fremde Hilfe angefertigt habe und nur die im Literaturverzeichnis angegebenen Quellen und Hilfsmittel benutzt habe.

_____, den _____