

Künstliche Intelligenz

ChatGPT: Eine KI, sie zu knechten

Seit Wochen staunen Laien über eine neue Künstliche Intelligenz (KI) namens „ChatGPT“, und die Fachleute wundern sich. Kann diese schlaue Frage-Antwort-Software zeigen, dass Maschinen irgendwann menschlich sein können?

Von Jörn Schumacher

5. Januar 2023



Foto: [Possessed Photography \(https://unsplash.com/photos/jlBMSMs4_kA\)](https://unsplash.com/photos/jlBMSMs4_kA)

Was sind die Chancen und Grenzen von Künstlicher Intelligenz? Darüber wird seit langer Zeit gestritten.

Die intelligente Software ChatGPT bringt Viele seit Wochen zum Staunen. Da stellt man in einem Chat-Fenster eine scheinbar schwierige Frage, und das Computersystem dahinter löst noch so komplizierte Aufgaben oder fasst lange Texte mit wenigen Sätzen zusammen. ChatGPT ist mit dem Wissen über die Welt gefüttert und in der Lage, Aussagen darüber zu machen; ChatGPT schreibt aber auch in Nullkommanichts auf Wunsch einen annehmbaren Haiku zu einem bestimmten Thema. Ebenso spuckt es nach kurzer Zeit den Quellcode für ein Computerprogramm aus, an dem ein Mensch eindeutig länger gesessen hätte. Von einem baldigen Ende des bisherigen Journalismus war schon die Rede.

Die Künstliche Intelligenz (KI), von der alle Welt spricht, stammt vom Unternehmen „OpenAI“, das vor sieben Jahren unter anderem vom Tesla-, Twitter- und SpaceX-Inhaber Elon Musk zusammen mit Microsoft gegründet wurde. Gerade für den Journalismus tut sich hier eine interessante Frage auf: Würde ein Leser einer KI als Autoren eines Artikels vertrauen? Tests mit ChatGPT haben längst gezeigt, dass seine Antworten alles andere als zuverlässig sind. Die Wissenschaftsjournalistin Eva Wolfangel befragte ChatGPT für einen Artikel in Spektrum der Wissenschaft nach

ihrem eigenen Namen und wer sich dahinter verbirgt. Was für einen Journalisten eine einfache Übung ist, brachte die viel gelobte KI-Software ins Schleudern. Einmal war Eva Wolfangel für die KI eine Cyber-Terroristin, kurz danach war sie eine Figur aus einem Science Fiction, und bei der dritten Anfrage spekulierte die KI wild eine Märchengeschichte zusammen.

Zusammengefasst: ChatGPT neigt zum Lügen. Das geben sogar die Entwickler selbst auf ihrer Webseite zu. Das Blöde ist nur: Man weiß als Anwender nicht, wann es lügt, und wann es die Wahrheit sagt. Für einen journalistischen Kontext das Todesurteil. Aber noch viel entscheidender: ChatGPT weiß selbst nicht, wann es lügt und wann nicht.

Moralischer Kompass eingebaut?

ChatGPT hat wie alle KI-Systeme kein Selbst-Bewusstsein. Es versucht nicht, wie die meisten Menschen, das eigene Tun mit einem moralischen Kompass möglichst überein zu bringen. Denn im normalen Leben steht man für eine Aussage, die man macht, mit seiner ganzen Person ein. Das heißt: Wenn ich lüge, identifizieren meine Gesprächsteilnehmer den Wahrheitsgehalt meiner Aussage mit meiner Person. Bin ich jemand, dem man vertrauen kann? Oder sage ich manchmal, aus unterschiedlichen Gründen, die Unwahrheit?

Jeder Gesprächsteilnehmer baut sich selbst in gewisser Weise ein Bild über das Gegenüber, eine Theorie darüber, wer der andere ist, welche Absichten er hat, welches Ziel er verfolgt. Und dazu gehört auch die Frage: Wie sieht sein moralischer Kompass aus? Muss ich damit rechnen, von ihm (sprachlich oder konkret) sozusagen übers Ohr gehauen zu werden? Damit verbunden ist ein Moral-System, von dem ich annehme, dass es in gewisser Weise jeder Mensch in sich trägt. Manches mag schräg oder im wahrsten Sinne des Wortes „verrückt“ sein, manches hoffnungslos verdorben; doch insgesamt sollten wir Menschen damit rechnen, dass andere Menschen irgendeine Art von Wertesystem aufgebaut haben.

Die Frage, ob ein Computer echte Intelligenz zeigt, hat beträchtliche soziale Auswirkungen, stellte bereits der Physiker Roger Penrose, einer der stärksten Kritiker einer „starken“ KI, in seinem Buch „Computerdenken“ fest. Ob eine KI irgendwann so sein könnte wie ein Mensch, ist eng verbunden mit dem sogenannten Leib-Seele-Problem, das in der Philosophie seit jeher heftig diskutiert wird und immer wieder in der Science-Fiction-Literatur durchgewalkt wird. Anhand des Androiden Data vom Raumschiff Enterprise etwa wurde so gut wie jedes erdenkliche Gedankenspiel zur Frage einmal durchgespielt. Etwa ob dieser Roboter dieselben Rechte hat wie ein Mensch, ob er einen freien Willen besitzt. Kurz: Ob eine Maschine einen Geist wie der Mensch besitzen kann? Eine Seele? Und was ist das überhaupt?

Das Interessante daran: Die Frage ist nie vollständig beantwortet worden. Die Diskussion hält an.

Vertreter der sogenannten „starken“ KI sind überzeugt, dass es lediglich eine Frage der Zeit und des technischen Aufwands ist, eine „ordentliche“ KI zu programmieren, die nicht nur menschlich wirkt, sondern menschlich ist. Das dürfte auch weiterhin so lange der Fall sein, wie man noch nicht definieren konnte, was die Seele eigentlich ist. Es reicht vielleicht gar nicht, eine Maschine mit ganz viel Wissen zu füttern sowie den Regeln, nach denen man dieses Wissen verknüpft, um dann Aussagen darüber in menschliche Sprache zu übersetzen.

Vielleicht bedarf es mehr. Kann sich zum Beispiel eine Maschine wundern? Oder sich ärgern? Wenn nein, wie sollte es diese Regungen dann im menschlichen Gegenüber jemals wirklich verstehen? Kann es moralische Ansprüche erheben, sich ungerecht behandelt fühlen, ein Wertesystem für sich beanspruchen und sein Recht einfordern?

Selbstbewusstsein? Fehlanzeige

Dafür bedarf es eines Bewusstseins von der Welt um sich herum und die eindeutige Abgrenzung seiner selbst von dieser Welt, sprich: ein Selbst-Bewusstsein. Nur wenn ich „Ich“ sagen kann und mich damit von allem, was um mich herum abgrenzen kann, kann ich auch für mich Dinge einfordern, etwas wünschen oder ersehnen. Erst dann kann ich ein Ziel verfolgen, um dieses Erwünschte zu erreichen, ich kann Schmerzen oder Schönheit empfinden, wenn mir ein Sehnen nicht erfüllt wird und so weiter.

Hinzu kommt die logische Folgerung daraus: Ein Ich, das „ich“ denken kann, kommt zu der Annahme, dass im Gegenüber, das menschlich zu sein scheint, ebenfalls ein Ich steckt. In der Psychologie und Philosophie spricht man von der „Theory of Mind“, ohne die menschliche Gegenüber immer nur Sachen – wie zum Beispiel ein Schrank – bleiben würden. Ich werde mein Gegenüber – halte ich es für einen Menschen – so behandeln, als hätten wir beide ein Wertesystem, sei es nun transzendent oder nicht.

„Der Mensch wird am Du zum Ich“, sagt Martin Buber. Kann ein Roboter, kann eine KI tatsächlich diese seelischen Zustände einprogrammiert bekommen – ohne dass er selbst wirklich gewahr werden kann, was es selbst ist? Gehört die Urfrage danach, woher wir kommen und wohin wir gehen, nicht zum Menschen wie seine DNA? Wie soll eine Maschine eine Antwort auf diese Frage jemals wirklich beantworten, ja: überhaupt erst richtig stellen können?

Kann eine Maschine, die auf bloßes Rechnen reduziert ist, wirklich immer

ausrechnen, was Wahrheit ist? Kann es die Frage nach seiner eigenen Existenz und nach deren Sinn in Wahrscheinlichkeitsparametern bestimmen? Oder kommt eine KI irgendwann an den Punkt, wo nur noch Glaube möglich ist? Seit den Erkenntnissen der Physik des 20. Jahrhunderts wissen wir, dass selbst die aufwändigste Rechenmaschine der Welt nicht den Lauf der Dinge berechnen kann. Ja, nicht einmal die Mathematik kann sichere Aussagen über alles treffen, wie der Mathematiker Kurt Gödel zeigte.

Die Frage geht sogar noch weiter. Am Ende steht hier die uralte Frage im Raum, was den Menschen überhaupt ausmacht und ob nur Materie ausreicht für sein Wesen, oder ob es da noch eines geistlichen, übersinnlichen, ja: göttlichen Zutat bedarf. Die Meinungen dazu teilten die Menschheit geradezu in zwei Lager, geschichtlich tauchen ihre Gegensätze immer wieder in Konflikten auf, und sie werden es wohl so lange tun, wie es Menschen gibt. Der Marxismus-Leninismus beispielsweise propagierte entschlossen einen reinen materiellen Realismus, demzufolge der Mensch nichts als Staub ist und ohne etwas Göttliches gedacht werden kann. Auch die darauf bauende Moral und politische Richtung muss folglich rein menschengemacht sein. Dem gegenüber steht ein dualistisch geprägtes Menschenbild (der Mensch ist Leib und Seele zugleich), das schon in der griechischen Philosophie seine Fundamente hat. Die Frage nach der Möglichkeit einer „starken“ KI führt also offenbar schnell zu den ganz großen Welt- und Menschheitsfragen, über die sich ganze Weltordnungen entzweit haben.

Für Christen ist es zudem ein größeres Problem, für einen noch so menschengleichen Roboter anzunehmen, dass er in den Heilsplan Gottes einen Platz haben könnte. Starb Jesus auch für eine KI? Kann eine KI sündigen? Steht der Name einer KI „geschrieben im Buch des Lebens vom Anfang der Welt an“ (Offenbarung 17,8)? Hat sie überhaupt einen freien Willen, sich für oder gegen etwas zu entscheiden?

Noch viel wichtiger als die Streitfrage, ob eine KI oder ein Roboter nun wirklich menschlich ist oder nicht, wird der Streit darüber werden, wie wir KIs und Roboter eines Tages behandeln wollen. Denn so uneindeutig diese Frage bislang entschieden ist, so vehement werden Vertreter der „starken“ KI dafür kämpfen, dass künstliche Wesen natürliche Rechte bekommen werden. Und was ist, wenn die Künstlichen Intelligenzen nicht mehr nur in einem Computer mit Bildschirm stecken, sondern eines Tages in einem menschenähnlichen Körper? Der Streit ist noch lange nicht zu Ende. Er hat noch nicht einmal richtig begonnen.

Helfen Sie PRO mit einer Spende

Bei PRO sind alle Artikel frei zugänglich und kostenlos - und das soll auch so bleiben. PRO finanziert sich durch freiwillige Spenden. **Unterstützen Sie**

jetzt PRO mit Ihrer Spende.

Jetzt spenden (<https://www.pro-medienmagazin.de/spenden/>)

Diese Website verwendet Akismet, um Spam zu reduzieren. Erfahren Sie mehr darüber, wie Ihre Kommentardaten verarbeitet werden (<https://akismet.com/privacy/>).

Link zum Beitrag:

<https://www.pro-medienmagazin.de/chatgpt-eine-ki-sie-zu-knechten/>