



[Kolumnen](#)[Reportagen](#)[Seiler kocht](#)[Planet Plüss](#)[Ausgaben](#)[Kreuzworträtsel](#)

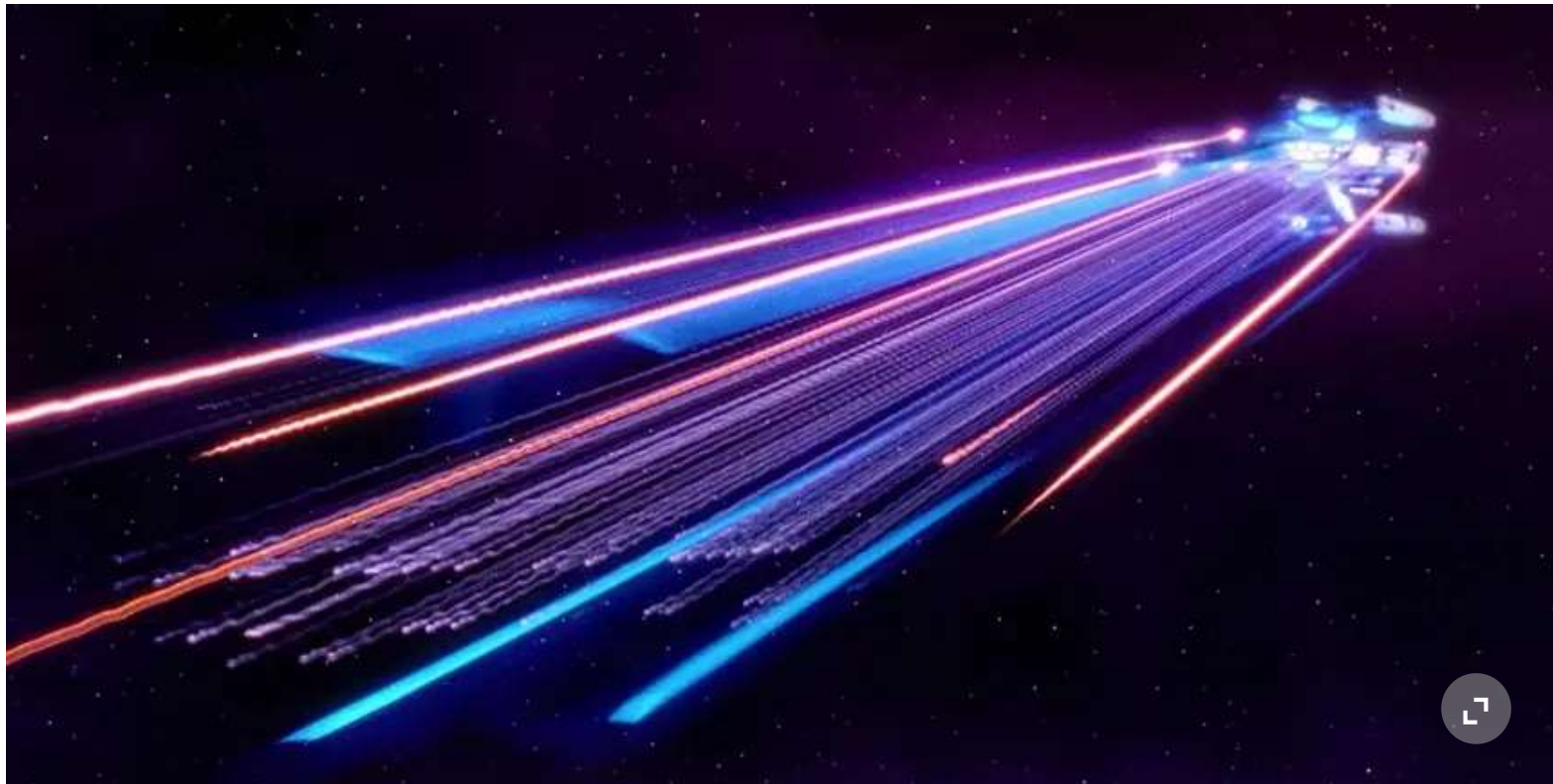
Abo [Ben Moore über KI in der Wissenschaft](#)

Was sind Raum und Zeit? Fragen Sie bloss nicht die KI

ChatGPT liest offenbar keine Bezahl-Zeitschriften, sagt der Astrophysiker. Und zeigt am Beispiel des Warp-Antriebs, warum bislang nur echte (und geniale) Menschen Antworten auf grosse Fragen geben können.

Ben Moore (Das Magazin)

Publiziert: 24.11.2023, 11:30



Warp-Antrieb ist wissenschaftlich unmöglich, behauptet die KI? Commander Spock vom Raumschiff Enterprise sollte ihr die Ohren lang ziehen.

Foto: Paramount Pictures

In meiner Anfangszeit als Astrophysiker lieferten unsere grössten Teleskope Bilder von Ausschnitten des Kosmos, die einige Tausend Galaxien enthielten. Ihre Formen und Grössen konnten von einem einzelnen Astronomen mit blossen Auge klassifiziert und katalogisiert werden, und es wurden auf diese Weise interessante Korrelationen zwischen verschiedenen Galaxientypen und ihrer Umgebung beobachtet. Etwa, dass rotierende Scheibengalaxien wie un-

sere Milchstrasse tendenziell allein oder in kleinen Gruppen vorkommen und nicht in grossen Galaxienhaufen.

Dies ist eine Beobachtung, die ich, wie viele andere Forscherinnen und Forscher auch, in meinen ersten Forschungsjahren zu erklären versucht habe, und die schliesslich in unser heutiges Verständnis der Entstehung von Galaxien mündete. Es stellte sich auch heraus, dass die Verteilung der Galaxien im Weltraum nicht zufällig ist – eine Entdeckung, die zur Entwicklung unseres Standardmodells der Kosmologie führte.

Heute, dreissig Jahre später, werden das Vera C. Rubin Observatory und das Weltraumteleskop Euclid bald Daten liefern, die nicht mehr einige Tausend, sondern Milliarden von Galaxien umfassen.

Diese riesigen Datensätze sind viel zu gross, als dass Menschen sie alle durchgehen könnten, um darin subtile Muster zu erkennen und Bezüge zu finden. Daher haben Astronomen für diese Aufgabe Software mit künstlicher Intelligenz entwickelt. Die Erwartungen an die Fähigkeiten der KI sind bekanntlich riesig. Aber was ist da dran – werden uns künftig Computer, und nicht mehr Menschen, die nächsten grossen Einblicke in die Geheimnisse unseres Universums ermöglichen?

Max übertreibt

Der jüngste KI-Hype dreht sich vor allem um ChatGPT, die bekannteste und vielleicht elaborierteste KI-Anwendung. Bill Gates erklärte, dass Werkzeuge wie ChatGPT eine ebenso wichtige Erfindung seien wie der PC und das Internet. Auf der anderen Seite sagte Sam Altman, der ehemalige CEO von OpenAI, dem Unternehmen, das ChatGPT entwickelt hat, dass künstliche Intelligenz wahrscheinlich zum Ende der Welt führen wird.

Über 30'000 KI-Forscherinnen und Branchenführer haben aufgrund der möglichen Gefahr eines Missbrauchs von KI einen Brief unterzeichnet, in dem sie eine sofortige Unterbrechung der Entwicklung von KI-Software fordern.

Diese Petition wurde vom Future of Life Institute formuliert, einer von Max Tegmark gegründeten Organisation. Ich war Anfang der Neunzigerjahre zur gleichen Zeit wie Max Postdoktorand im Astronomie-Department der Universität Berkeley, und wir führten viele Gespräche beim Feierabendbier auf Berkeleys berühmter Telegraph Avenue, der Keimzelle der amerikanischen Hippie-Bewegung und Counterculture der Sechzigerjahre, die nicht nur die USA, sondern die ganze Welt verändert hatten.

**Trotz der Entwicklung
immer ausgefeilterer
Chatbots bestand kein
einziges den Turing-Test
auch nur annähernd. Bis
ChatGPT kam.**

Max ist ein kluger Kerl und hat tiefe Erkenntnisse über die mathematische Natur unseres Universums gewonnen und publiziert. Aber stellen ChatGPT und ähnliche Anwendungen tatsächlich eine Gefahr für die Menschheit dar, oder können wir sie nutzen, um die Natur von Raum, Zeit und Schwerkraft zu verstehen?

Im Jahr 1950 fragte sich der Begründer der künstlichen Intelligenz, der englische Mathematiker Alan Turing, ob ein Computerprogramm über eine Textschnittstelle mit einem Menschen kommunizieren und ihn derart täuschen könnte, dass dieser meint, sich mit einem anderen Menschen zu unterhalten.

Ein solches Programm wird Chatbot genannt, und die Frage ist inzwischen als Turing-Test der künstlichen Intelligenz bekannt. Der erste Chatbot namens ELIZA wurde 1966 programmiert, und trotz der Entwicklung immer ausgefeilterer Chatbots bestand kein einziger den Turing-Test auch nur annähernd. Bis ChatGPT kam.

Hausarbeit und Quantensprung

Ende letzten Jahres wurde ChatGPT-3, die im Vergleich zu ihren beiden Vorgängerinnen deutlich verbesserte Version, auf die Welt losgelassen, und sie war ein Quantensprung nach vorne. Frühere Chatbots konnten ein oder zwei sinnvolle Sätze produzieren, wenn ihnen eine Frage gestellt wurde, kamen dann aber schnell vom Thema ab und gaben Unsinn von sich.

Richtet man dagegen an ChatGPT eine detaillierte Aufforderung, zum Beispiel, einen Aufsatz zu einem bestimmten Thema zu schreiben, kommt dabei ein Text heraus, der nachweislich eine bessere Note erhält als einer, den ein durchschnittlicher Universitätsstudent selbst schreibt.

Ich befürchte, dass einige der Hausarbeiten, die meine Studentinnen und Studenten im letzten Semester eingereicht haben, mithilfe

von ChatGPT geschrieben wurden, aber das lässt sich leider nicht beweisen.

**Fragt man ChatGPT nach
einem Rezept für etwas,
das es nicht gibt, zum
Beispiel eine Schokoladen-
kohlsuppe, so erhält man
eine ausführliche
Anleitung.**

Kein derzeit vorhandenes Softwaretool kann wirklich unterscheiden, ob ein Text von einem Menschen oder von einem Computer geschrieben wurde. Wir müssen darüber nachdenken, wie wir Studenten dazu ermutigen können, solche Software zu nutzen, um besser zu lernen, und nicht als einfache Möglichkeit zur Erledigung von Hausarbeiten. Aber das ist nur eines von mehreren Missverständnissen.

ChatGPT, auch das denken viele, ist kein «besseres» Google – es durchsucht nicht einfach das Internet nach Antworten. Fragt man ChatGPT nach einem Rezept für etwas, das es nicht gibt, zum Beispiel eine Schokoladenkohlsuppe, so erhält man eine ausführliche Anleitung mit den entsprechenden Warnungen zur Verträglichkeit.

ChatGPT kann Computercode schreiben und mit einer Genauigkeit von über 75 Prozent klinische Diagnosen erstellen. Es kann sogar Gedichte schreiben, wenn auch keine guten. Das ist wiederum gut für uns, denn im Idealfall sollte die KI repetitive, langweilige Arbeit automatisieren, die niemand mag.

Sie sollte die riesigen Datensätze, die wir generieren, analysieren und so uns Menschen mehr Zeit geben, kreativ zu sein.



Das Rubin Observatorium in Chile ist noch im Bau: Dank KI wird der Mensch von hier aus die Milchstrasse noch gründlicher durchforsten können.

Foto: Bruno C. Quint

Bei ChatGPT kommt ein komplexes statistisches Modell zur Anwendung, das vorhersagt, welche Wörter in einer Sequenz als nächste

kommen. Das Modell wird mithilfe eines künstlichen neuronalen Netzwerks erstellt, das der Funktionsweise unseres Gehirns ähneln soll. Es funktioniert auf die gleiche Weise wie die KI-Werkzeuge, die Astronomen entwickeln – indem Eingabedaten über miteinander verbundene Schichten künstlicher Neuronen verarbeitet werden.

Wenn diese Programme trainiert werden, «lernen» sie, die Stärke der Verbindungen anzupassen, um Muster und Beziehungen in den Daten zu erkennen. Diese gelernten Informationen werden dann verwendet, um Vorhersagen zu treffen oder Aufgaben auszuführen.

Peinliche Fehler

Während unsere Software auf astronomischen Daten basiert, hat ChatGPT das gesamte Internet als Lernstoff. Das sind etwa eine Billion Wörter auf mehreren Milliarden von Menschen geschriebenen Websites, einschliesslich der riesigen Datenbanken von Konversationen und Kommentaren auf Websites wie Reddit oder auf Blogs. Zudem stehen ChatGPT etwa fünf Millionen digitalisierte Bücher (von etwa 100 Millionen Büchern, die jemals geschrieben wurden) zur Verfügung, was weitere 100 Milliarden Wörter potenzielles Lernmaterial bietet.

Niemand versteht dabei die detaillierte Funktionsweise des Sprachmodells von ChatGPT, das über eine Billion Parameter umfasst. Sogar die Programmierer, die es entwickelt haben, waren von seiner Leistungsfähigkeit überrascht. Hunderte von Firmen haben ChatGPT bereits in ihr Unternehmen integriert und den Dienst kostenpflichtig abonniert.

Allerdings sind die Autorinnen und Autoren, welche die Inhalte erstellt haben, anhand derer das Programm geschult wurde, zu Recht verärgert. Ihre Werke und Texte – alles, was jemals einer von uns online von sich gegeben hat – werden ohne Genehmigung verwendet, und niemand hat dafür eine Vergütung erhalten. Wahrscheinlich wurden auch einige meiner Bücher und Kolumnen ungefragt übernommen, was mich ziemlich irritiert.

**Ich halte es für unmöglich,
dass KI etwas so Tief-
gründiges wie das Rad
erfinden könnte – etwas zu
konzipieren, das nicht
existiert, ist eine einzig-**

artige Fähigkeit des Menschen.

Neben der fragwürdigen Weise, in der das Programm seine Aufgabe «gelernt» hat, zeigt ChatGPT auch inhaltlich deutliche Mängel. Es produziert Texte, die nicht sonderlich kreativ sind, also so wie die durchschnittliche Textqualität, die man im Internet findet.

In den seltenen Fällen, in denen es Stellung bezieht, ist es frustrierend «moralisch», und seine Antworten sind dank handcodierter Regeln seiner Software-Ingenieure ethisch hochgradig bereinigt. Gleichzeitig formuliert es seine Aussagen überaus selbstbewusst, selbst dann, wenn sie sachlich falsch sind.

Wie gefährlich solche Falschaussagen sind, wurde für die Muttergesellschaft von Google erschreckend deutlich, die 100 Milliarden Dollar an Aktienwert verlor, nachdem ihre KI-Anwendung Bard – ein Konkurrenzprodukt zu ChatGPT – gleich bei seiner ersten Demonstration einen ziemlich peinlichen Fehler machte. Bard wurde der Welt Anfang dieses Jahres mit der Frage vorgestellt: «Von welchen neuen Entdeckungen des James-Webb-Weltraumteleskops kann ich meinem neunjährigen Kind erzählen?»

Bard gab an, dass das Teleskop das allererste Bild eines Exoplaneten gemacht habe, also eines Planeten, der einen Stern ausserhalb unseres Sonnensystems umkreist. Tatsächlich wurde das erste derartige Bild aber bereits 2004 vom Very Large Telescope des European Southern Observatory aufgenommen.

KI-Programme werden sicherlich in vielen Berufen die Produktivität steigern und in einigen den Menschen vollständig ersetzen. Die eigentliche Gefahr besteht jedoch darin, dass wir uns zu sehr auf sie verlassen und sie kriminell oder in schlechter Absicht eingesetzt werden, von der Entwicklung biologischer Krankheitserreger bis hin zu Cyberangriffen oder der Verbreitung von Desinformation.

Die Tech-Milliardäre, die in diese Softwaretools investieren, konzentrieren sich auf das zukünftige Risiko, dass KI hyperintelligent wird, ähnlich wie Skynet aus den «Terminator»-Filmen – ein Gruppengeist von Computern und Maschinen mit eigenem Bewusstsein, der versucht, die Menschheit auszulöschen, weil er uns als Bedrohung wahrnimmt. Dieser Fokus auf die ferne Zukunft ist eine clevere Taktik, um von den aktuellen Gefahren abzulenken, die diese Investoren möglicherweise verursacht haben.

Sinn und Systemgrenze

Ich stimme dieser dystopischen Sichtweise ohnehin nicht zu. Die KI-Programme sind nicht das, was ich als intelligent bezeichnen würde. Ich glaube nicht, dass sie mit ihrer derzeitigen Arbeitsweise die menschliche Intelligenz übertreffen werden, weil sie nicht wie ein Mensch lernen.

Ausser mit Sprache und Geräuschen wird unser Gehirn auch mit Sehen, Berühren, Riechen und mehr als einem Dutzend anderer Sinnesreize trainiert. Wie kann KI kreative Prosa über den Duft eines Apfels schreiben, den sie noch nie gerochen hat, oder über die Magie eines Sonnenuntergangs, den sie nie gesehen hat? Ich halte es auch für unmöglich, dass KI etwas so Tiefgründiges wie das Rad erfinden könnte – etwas zu konzipieren, das nicht existiert, ist eine einzigartige Fähigkeit des Menschen.

**Es macht keinen Sinn, die
KI zu fragen: «Was ist
Raum, was ist Zeit?» Sie
wird immer die gleichen
langweiligen Antworten
geben.**

Ich bezweifle auch, dass KI es uns in naher Zukunft ermöglichen wird, die Geheimnisse unseres Universums zu lüften. Neulich fragte ich ChatGPT, wie ich ein Raumschiff mit Warp-Antrieb bauen könnte, also schneller als das Licht, so wie wir es aus «Star Trek» kennen. ChatGPT antwortete, dass dies unmöglich sei und den Rahmen der theoretischen Physik sprengte.

Offensichtlich gehörte die kostenpflichtige Fachzeitschrift «Classical and Quantum Gravity» nicht zum Lernstoff von ChatGPT. In einem Essay der Zeitschrift wurde ein funktionierender Warp-Antrieb ausführlich diskutiert und tatsächlich für möglich erachtet, wenn er auch in der Praxis nur sehr schwer umsetzbar wäre.

Es macht keinen Sinn, ChatGPT zu fragen: «Was ist Raum, was ist Zeit?» Oder es aufzufordern: «Entwickle eine neue Strategie, um Einsteins Theorien zu testen.» Es wird immer die gleichen langweiligen und uninspirierten Antworten geben. Und auch jenseits von ChatGPT lassen die Segnungen der KI für die Astronomie noch auf sich warten: Ein von Astronomen konzipiertes und trainiertes KI-Werkzeug durchforstet zwar unermüdlich unsere grossen astronomischen Datensätze, tiefgreifende Erkenntnisse hat es daraus bisher aber nicht gewonnen.

Ben Moore ist Professor für Astrophysik an der Universität Zürich.
redaktion@dasmagazin.ch

Fehler gefunden? [Jetzt melden.](#)

6 Kommentare