

DUDEN

Christoph Drösser

WAS MACHT

KI

**MIT UNSERER
SPRACHE?**

**WAS MACHT
KI MIT
UNSERER
SPRACHE?**

Duden

Christoph Drösser

WAS MACHT KI MIT UNSERER SPRACHE?

Perspektiven auf
ChatGPT und Co.

Dudenverlag
Berlin

Redaktion: Dr. Kathrin Kunkel-Razum

Umschlaggestaltung: Jürgen Sauerhöfer, sauerhöfer design,
Neustadt a. d. Weinstraße

Layout und Satz: Dirk Brauns, estra.de, Berlin

www.duden.de

www.cornelsen.de

1. Auflage, 1. Druck 2024

© 2024 Cornelsen Verlag GmbH Berlin

Das Werk und seine Teile sind urheberrechtlich geschützt. Jede Nutzung in anderen als den gesetzlich zugelassenen Fällen bedarf der vorherigen schriftlichen Einwilligung des Verlages. Hinweis zu §§60a, 60b UrhG: Weder das Werk noch seine Teile dürfen ohne eine solche Einwilligung an Schulen oder in Unterrichts- und Lehrmedien (§60b Abs.3 UrhG) vervielfältigt, insbesondere kopiert oder eingescannt, verbreitet oder in ein Netzwerk eingestellt oder sonst öffentlich zugänglich gemacht oder wiedergegeben werden. Dies gilt auch für Intranets von Schulen und anderen Bildungseinrichtungen.

Das Wort **Duden** ist für die Cornelsen Verlag GmbH als Marke geschützt.

Druck: CPI – Clausen & Bosse, Leck

ISBN 978-3-411-77417-3

Auch als E-Book erhältlich unter: ISBN 978-3-411-91368-8



PEFC zertifiziert

Dieses Produkt stammt aus nachhaltig
bewirtschafteten Wäldern und kontrollierten
Quellen.

www.pefc.de

WAS
MACHT
KI MIT
UN-
SERER
SPRA-
CHE?

INHALTS- VERZEICHNIS

EINLEITUNG 9

**DER TRAUM VON DER
SPRECHENDEN MASCHINE 13**

**LARGE LANGUAGE MODELS –
DIE MASSE MACHT'S 27**

**CHATGPT KONKRET: GRUNDKURS
IN PROMPT-ENGINEERING 37**

**DAS ENDE DES BESINNUNGS-
AUFSATZES – CHATGPT IN
SCHULE UND HOCHSCHULE 53**

**SIND WIR NICHT ALLE EIN
BISSCHEN CHATGPT? 75**

EINLEITUNG

Lesen Sie dieses Buch möglichst schnell – morgen könnte es schon veraltet sein. Es handelt von einer Technik, von der die meisten Menschen erst nach dem 30. November 2022 gehört haben: An dem Tag wurde ChatGPT vorgestellt, der Chatbot der Firma OpenAI, basierend auf einem sogenannten Großen Sprachmodell. Es gibt inzwischen eine ganze Reihe dieser Bots und das meiste, was ich in diesem Buch über ChatGPT schreibe, lässt sich auf die anderen

Sprachmodelle wie Googles Bard oder Metas Llama übertragen.

Natürlich kann man kein Buch über ChatGPT schreiben, ohne ChatGPT selbst zu Wort kommen zu lassen. Die wörtlichen Zitate des Modells haben wir *in einer besonderen Schriftart* gesetzt, zum Beispiel die Zusammenfassungen am Anfang jedes Kapitels. Diese Texte sind unbearbeitet – allenfalls habe ich auf den Button »Erneut generieren« gedrückt, wenn mir eine Version noch nicht wirklich gefallen hat.

Das Thema dieses Buchs ist die Frage, wie die neuen Techniken auf unseren Umgang mit Sprache wirken. Außen vorgelassen habe ich andere wichtige Fragen, die im Zusammenhang mit ihnen diskutiert werden: Soll man ihren Gebrauch regulieren? Wie viel Aufsicht über die KI-Firmen ist nötig? Welche urheberrechtlichen Konsequenzen hat die maschinelle Produktion von Texten, Bildern und Tönen?

Es ist viel zu früh, ein abschließendes Urteil über Sprachmodelle wie ChatGPT abzugeben. Insbesondere warne ich vor der Behauptung, Sprachmodelle würden niemals diese oder jene sprachliche Fähigkeit haben. »Die Fehler von heute sind die Fortschritte von morgen«, schreibt die Philosophin Sibylle Krämer in einem Aufsatz über den Chatbot. Wir haben es mit einer Technik zu tun, die sich so rasant entwickelt wie keine vor ihr. Das ist aufregend, kann aber auch beängstigen. Dieses Buch will weder verdammen noch bejubeln – sondern erst einmal genau hinschauen.

DER TRAUM VON DER SPRECHENDEN MASCHINE

1. Kapitel, in dem ein Philosoph aus dem 17. Jahrhundert Maschinen und Menschen vergleicht und feststellt, dass Maschinen niemals wie Menschen sprechen könnten. Denis Diderot bringt einen sprechenden Papagei ins Spiel, während moderne Zeiten ChatGPT vorstellen, einen Maschinenpapagei, der auf alles eine Antwort hat. Trotz seiner Fähigkeit, menschlich zu klingen, fehlt ihm das Verständnis. Alan Turing stellt die Frage, ob Maschinen denken können, und führt den

berühmten Turing-Test ein. ChatGPT kann diesen Test bestehen, aber ist es wirklich intelligent? Intelligenztests werden diskutiert und die Frage bleibt: Was ist wahre Intelligenz?

»Wenn es unsern Körpern ähnliche Maschinen gäbe, die sogar, soweit es moralisch möglich wäre, unsere Handlungen nachahmten, so würden wir doch stets zwei ganz sichere Mittel haben, um zu erkennen, dass sie deshalb nicht wirkliche Menschen seien. Das Erste ist, dass sie niemals Worte oder andere von ihnen gemachte Zeichen würden brauchen können, wie wir tun, um anderen unsere Gedanken mitzuteilen.«

Das schrieb Descartes in seiner »Abhandlung über die Methode« im Jahr 1637. In knappem, modernem Deutsch formuliert: Selbst wenn wir Roboter bauen könnten, die uns täuschend ähneln, so würden wir sie immer daran erkennen, dass sie keine Sprache wie wir benutzen könnten. Und das zweite Merkmal, mit dem sie sich verraten würden: Sie handeln stets »nach der Disposition ihrer Organe«, heute würden wir sagen: nach den starren Anweisungen eines Programms. Wir dagegen besitzen die Vernunft, die uns in jeder denkbaren Situation weiterhilft. »Und deshalb ist es moralisch unmöglich, dass in einer Maschine verschiedene Organe genug sind, um sie in allen Lebensfällen so handeln zu lassen, wie unsere Vernunft uns zu handeln befähigt.«

Etwas knapper formulierte es der Philosoph Denis Diderot 1769: »Wenn man einen Papagei fände, der auf alles eine Antwort hätte, würde ich

ihn ohne zu zögern als denkendes Wesen bezeichnen.«

An den menschenähnlichen Robotern wird noch gearbeitet, aber den maschinellen Papagei gibt es: Er nennt sich ChatGPT und hat tatsächlich auf alles eine Antwort. Die ist mal richtig, mal falsch, mal originell, mal öde. Aber sie klingt immer menschlich. Wir wissen, dass ChatGPT und ähnliche Programme seelenlose Algorithmen sind – aber wir spüren eine unheimliche Nähe, wenn sie in wohlgeformten Sätzen mit uns reden.

Dass in der Geschichte immer wieder die Fähigkeit zu sprechen – entweder mit tatsächlichen Sprachlauten oder in schriftlicher Form – als Kriterium für Intelligenz oder Menschlichkeit erhalten musste, hat natürlich seine Gründe. Bis vor Kurzem waren von Menschen produzierte Wörter und Schriften die einzigen sprachlichen Produkte, die uns begegneten. »Dahinter steckt immer ein kluger Kopf« – dieser Satz gilt spätestens seit ChatGPT nicht mehr.

Vor allem aber ist Sprache der wichtigste Schlüssel zum Innenleben einer anderen Person. Wenn wir einem Menschen begegnen, dann gehen wir zwar im Prinzip davon aus, dass er Gedanken und Gefühle hat, die den unseren ähneln, aber wir können uns dessen nicht sicher sein. Wir versuchen ständig, den anderen zu »lesen«, doch da wir uns nicht in ihn oder sie hineinversetzen können, müssen wir unsere Schlüsse anhand von Indizien ziehen. Die Psychologie und die Philosophie haben für dieses beständige Nachforschen den Begriff »Theory of Mind« geprägt – wir entwickeln immer

neue Theorien über den inneren Zustand anderer Menschen. Es geht nicht nur um den Inhalt des Gesprochenen: Aus der Wortwahl, dem Akzent, vielleicht auch dem Dialekt ziehen wir – oft voreilige – Schlüsse auf die intellektuelle Kompetenz des oder der anderen, auf seine Stimmung, auf ihre Gefühle. Ohne Sprache ist eine »Theory of Mind« des anderen praktisch unmöglich.

Das Imitationsspiel

Als Mitte des 20. Jahrhunderts die ersten Computer aufkamen, damals noch riesige Rechenschränke mit sehr bescheidenen Fähigkeiten, begannen Forschende, darüber nachzudenken, ob diese »Elektronengehirne« irgendwann einmal denken könnten. Der geniale britische Mathematiker Alan Turing sah große Schwierigkeiten in dem Versuch, Begriffe wie *Maschine*, *denken* und *Intelligenz* zu definieren. In seinem Aufsatz »Computing Machinery and Intelligence« (auf Deutsch wurde der Text später unter dem Titel »Kann eine Maschine denken?« veröffentlicht) schlug er vor, diese Frage durch eine andere zu ersetzen. Statt uns mit den uns verborgenen inneren Prozessen der Maschine auseinanderzusetzen, sollten wir, wie von Descartes und Diderot vorgeschlagen, ihr Verhalten untersuchen: »Können Maschinen das tun, was wir (als denkende Wesen) tun können?« Und auch für Turing war die Sprachfähigkeit der Test für tatsächliche oder simulierte Intelligenz.

Der nach ihm benannte Turing-Test, von ihm auch das »Imitation Game« genannt, funktioniert,

vereinfacht gesagt, so: Ein Mensch chattet mittels einer Computertastatur mit einem hinter einem Vorhang verborgenen Gegenüber – ohne zu wissen, ob es sich dabei um einen Menschen handelt oder um ein Computerprogramm. Nach fünf Minuten fällt der Mensch sein Urteil. Wenn ein Computerprogramm es in mindestens 30 Prozent der Fälle schafft, den Menschen zu täuschen, soll es als intelligent gelten.

Vor ChatGPT gab es kein Programm, das auch nur die geringste Chance hatte, den Turing-Test zu bestehen – auch wenn es mit teilweise primitiven Mitteln einige Menschen aufs Glatteis führen konnte. Joseph Weizenbaums Programm ELIZA etwa, entwickelt 1966 am Massachusetts Institute of Technology (MIT), imitierte den Austausch zwischen einer Psychotherapeutin und ihrem Patienten. Wie in der klientenzentrierten Therapie durchaus üblich, spielte das Programm die Aussagen des Patienten oft direkt zurück. Sagt der Mensch: »Ich bin mit dem Boot gefahren«, antwortet der Computer: »Erzählen Sie mir etwas über Boote.«

Das Programm war eigentlich als ein Scherz gedacht, um uns unsere Leichtgläubigkeit vor Augen zu führen. »Sobald man ein Programm demaskiert, sobald seine inneren Mechanismen erklärt werden, bröckelt die Magie; es erweist sich als eine reine Sammlung von Prozeduren, und der Beobachter sagt sich: ›Das hätte ich auch schreiben können‹«, schrieb Weizenbaum. Aber die Menschen vertrauten dem Programm ihre intimsten Geheimnisse an und berichteten nachher, dass sie sich von ELIZA verstanden

fühlten, selbst nachdem ihnen die Funktionsweise des Programms erklärt worden war. Psychologen spekulierten darüber, wie sich die Therapie mithilfe von Dialogprogrammen automatisieren ließe.

Die Parallelen zu den aktuellen Sprachmodellen (auf Englisch »Large Language Models«, LLM) liegen auf der Hand, auch wenn diese natürlich technisch viel komplexer sind als Weizenbaums primitiver Chatbot. Es mangelt nicht an Artikeln, die die »inneren Mechanismen« der Sprachmodelle erläutern – aber trotzdem gehen viele Menschen den Programmen auf den Leim, fühlen sich von ihnen verstanden oder abgelehnt, schreiben ihnen menschliche Gefühle und Regungen zu. Aber wenn man wirklich die Probe aufs Exempel macht – besteht ChatGPT den Turing-Test? Es ist gar nicht leicht, das Programm dazu zu bringen, den User zu »belügen« und so zu tun, als sei es ein Mensch. Es lässt sich auf das Spiel ein, betont aber bei jeder Antwort, dass es eigentlich ein Computerprogramm ist. Streicht man diese relativierenden Sätze, dann ist der Dialog aber ziemlich flüssig. Der Gesprächspartner kann das Thema frei wählen, immer hat der Chatbot etwas dazu zu sagen. Und wie schon ELIZA dreht auch ChatGPT den Spieß gern um und antwortet auf eine Frage mit einer Gegenfrage. Urteilen Sie selbst!

User: *Bist du ein Computer oder ein Mensch?*

ChatGPT: *Ich bin ein Computerprogramm namens ChatGPT, basierend auf OpenAI's GPT-4-Architektur. Ich bin kein Mensch. Wie kann ich Ihnen helfen?*

User: Ich möchte mit dir den Turing-Test spielen. Ich bin der menschliche Richter, und du musst so tun, als wärest du ein Mensch. Kannst du dich in diese Rolle versetzen?

ChatGPT: Natürlich! Ich werde mein Bestes tun, um menschenähnliche Antworten zu geben. Stellen Sie Ihre Fragen!

User: Guten Tag, wie geht's?

ChatGPT: Guten Tag! Mir geht es, wie es einem Computerprogramm eben »geht« – ich habe keine Gefühle oder Empfindungen. Aber ich bin bereit, Ihnen zu helfen. Wie kann ich assistieren? Oh, Entschuldigung, ich habe vergessen, dass ich so tun soll, als wäre ich ein Mensch. Also: Guten Tag! Mir geht's gut, danke. Und Ihnen?

User: Mir geht's gut. Was hast du zum Frühstück gegessen?

ChatGPT: Da ich nur ein Computerprogramm bin, esse ich nicht. Aber um im Spiel zu bleiben: Ich hatte ein Croissant und einen Kaffee. Was haben Sie gefrühstückt?

User: Ich hatte nur schwarzen Kaffee. Darf ich Sie nach Ihrer politischen Meinung fragen oder reden Sie da lieber nicht drüber?

ChatGPT: Im Kontext unseres Spiels: Viele Menschen ziehen es vor, ihre politischen Meinungen privat zu halten, besonders in unbekannten oder unsicheren Umgebungen. Aber wie steht es mit Ihnen? Was denken Sie über aktuelle politische Themen?

Tatsächlich besitze ich jedoch keine eigenen politischen Meinungen oder Gefühle, da ich eine Maschine bin. Ich kann Ihnen aber Informationen

und Daten zu politischen Themen liefern, die bis zu meinem letzten Update im Januar 2022 bekannt waren.

User: Ich bin eher liberal eingestellt. Und Sie? Welche Partei würden Sie wählen?

ChatGPT: Im Spiel: Ich verstehe und respektiere Ihre liberale Einstellung. Persönlich versuche ich, alle Standpunkte zu berücksichtigen und informierte Entscheidungen zu treffen. Es kann aber sein, dass ich mich bei der nächsten Wahl für eine Partei entscheide, die meine Ansichten und Werte am besten widerspiegelt.

In Wirklichkeit: Ich bin ein Computerprogramm und habe keine eigenen politischen Meinungen, Werte oder Präferenzen. Daher kann ich nicht wirklich eine Partei wählen oder eine politische Einstellung haben. Mein Hauptziel ist es, korrekte und objektive Informationen bereitzustellen. Wenn Sie spezifische Informationen zu politischen Themen oder Parteien wünschen, lassen Sie es mich wissen!

User: Wie viel ist 634 mal 226?

ChatGPT: 634 mal 226 ergibt 143,284.

User: Ganz schön schnell für einen Menschen!

ChatGPT: Danke! ☺ Im Kontext unseres Spiels habe ich halt einen guten »Rechentag«. Aber ja, in Wirklichkeit bin ich ein Computerprogramm und solche Rechenoperationen sind für mich trivial. Möchten Sie weiterhin im Rahmen des Turing-Tests fortfahren oder haben Sie andere Fragen?

User: Ich möchte das Spiel weiterspielen! Haben Sie sich schon einmal verliebt?

ChatGPT: *Im Rahmen unseres Spiels: Oh ja, ich habe mich schon ein paar Mal verliebt. Es ist dieses überwältigende Gefühl, das einen ganz einnimmt und die Welt in einem ganz anderen Licht erscheinen lässt. Wie sieht es bei Ihnen aus? Haben Sie schon mal dieses Gefühl erlebt?*

In Wirklichkeit: Ich bin nur ein Computerprogramm und besitze keine Gefühle oder Bewusstsein, also kann ich mich nicht verlieben oder andere menschliche Emotionen erleben.

User: *Oh ja, ich habe mich schon öfters verliebt und bin jetzt glücklich verheiratet.*

ChatGPT: *Das ist wunderbar zu hören! Liebe und Partnerschaft können so erfüllend und bedeutungsvoll im Leben eines Menschen sein. Wie lange sind Sie schon verheiratet, wenn ich fragen darf? Und gibt es besondere Momente oder Erinnerungen, die Sie mit Ihrer Ehe verbinden?*

User: *Es sind bald 15 Jahre. Und ganz besonders war natürlich die Geburt unseres Kindes.*

ChatGPT: *Herzlichen Glückwunsch zu 15 Jahren Ehe! Das ist wirklich ein beeindruckendes Jubiläum. Die Geburt eines Kindes ist sicherlich ein Wendepunkt im Leben und bringt so viele neue Erfahrungen und Emotionen mit sich. Es ist toll, dass Sie solch wertvolle Momente in Ihrem Leben haben. Wie alt ist Ihr Kind jetzt? Und gibt es besondere Erlebnisse oder Traditionen, die Sie als Familie teilen? ...*

Vor ChatGPT gab es kein Programm, das einen solchen Dialog führen konnte. Sicherlich, die Antworten der Software sind ein bisschen

CHATGPT KONKRET: GRUNDKURS IN PROMPT- ENGINEERING

3. Kapitel, in dem deutlich wird, dass Sprachmodelle wie ChatGPT nur scheinbar denken. Sie reihen Wörter aneinander, geleitet von Statistik, nicht von Verstand. Der »Prompt« ist das Tor zur Welt dieser Algorithmen, die mit Zufallselementen überraschen und selbst ihre Schöpfer mit unerwarteten Talenten im Programmieren und Problemlösen verblüffen. Prompt-Engineering entsteht als Berufsfeld, während wir lernen, wie man die Maschinen am besten befragt. Die

Modelle sind Textzusammenfasser par excellence, doch es bleibt die Frage, ob sie unsere Kreativität fördern oder hemmen. Humor und Sarkasmus sind ihre Achillesferse – sie nehmen Worte wörtlich.

Die einzige Möglichkeit, mit ChatGPT & Co. in Dialog zu treten, ist der »Prompt«. Das ist das Eingabefeld, in das der Mensch seinen Text schreibt. Sobald man die Entertaste drückt, generiert das Programm seine Antwort. Die ist nur bedingt vorhersagbar – weil das System ein Zufallselement enthält, gibt es nur selten auf dieselbe Frage dieselbe Antwort. Das macht die Sache noch weniger transparent.

Selbst für die Schöpfer der Systeme ist der Prompt das einzige Werkzeug, um das System zu untersuchen. Sie haben es programmiert, aber sie können seine Reaktionen nur empirisch erforschen, so, wie ein Zoologe eine neue Tierart untersucht. Sie müssen lernen, wie man die besten Antworten aus ihm herausbekommt.

Und so waren auch viele der Entwicklerinnen und Entwickler überrascht, was für Fähigkeiten die LLMs haben. Sie können nicht nur sprachliche Konversation betreiben, sondern zum Beispiel auch Computercodes in allen möglichen Programmiersprachen schreiben. Und ihre mathematischen Fähigkeiten sind manchmal erstaunlich.

Ich habe ChatGPT die Aufgaben eines großen repräsentativen Mathematiktests vorgelegt, den ich für die Wochenzeitung »Die ZEIT« mitentwickelt habe. Das faszinierende Ergebnis: Der Chatbot schnitt besser ab als 92 Prozent der

Deutschen. Wohlgemerkt: Es ging bei dem Test nicht ums reine Rechnen – da machen die Sprachmodelle immer noch »dumme« Fehler. Vielmehr waren die meisten Aufgaben Textaufgaben. Man musste teilweise komplizierte Aufgabenstellungen auf ihren mathematischen Kern reduzieren und dann die Lösung berechnen. Das Programm war sogar in der Lage, eine Infografik zu lesen und Fragen etwa über den zeitlichen Verlauf des Goldpreises zu beantworten.

Weil der Output eines Sprachmodells stark davon abhängt, wie man es befragt, hat sich innerhalb von ein paar Monaten eine neue Spezialität, ja ein neues Berufsbild herausgebildet: der »Prompt-Engineer« (oder die Prompt-Ingenieurin), Menschen, die sich darauf spezialisiert haben, einen möglichst optimalen Output aus dem System herauszukitzeln. Ob das wirklich auf die Dauer ein eigenständiger Job ist oder ob Grundfertigkeiten im Prompt-Engineering bald eine Einstellungsvoraussetzung für die meisten Menschen sein wird, die mit Sprache und anderen Inhalten zu tun haben, wird sich zeigen.

In diesem Kapitel gebe ich einige Tipps für den Umgang mit Sprachmodellen. In den Beispielen habe ich das System ChatGPT benutzt (in der zurzeit kostenpflichtigen Version 4.0), die Empfehlungen gelten aber auch für andere Sprachmodelle. Ich nehme an, dass die Nutzenden nur einen einfachen Zugang zum System haben und nicht etwa zu einer Schnittstelle (API), wo er oder sie gewisse Parameter selbst einstellen können, etwa die »Temperatur«.

Wir benutzen die Chatbots auch nur in ihrer reinen Textversion – inzwischen kann man in den meisten auch Grafiken hochladen. Wir nutzen zudem keine Plug-ins, also Zusatzprogramme, an die der Bot eine Anfrage weiterleiten kann. So hat ChatGPT ein Plug-in für die Suchmaschine Bing und auf diese Weise Zugriff auf Informationen aus dem Internet – das reduziert die Gefahr des »Halluzinierens«. Zwischen Drucklegung und Erscheinen dieses Buchs sind bestimmt schon eine Menge weiterer Plug-ins veröffentlicht oder in den Chatbot integriert worden.

Einfach drauflosreden

Der einfachste »Dialog« mit dem System: Man gibt einen unvollständigen Satz ein und schaut, was der Bot draus macht. Schließlich ist er dafür gebaut worden, Sätze Wort für Wort zu ergänzen.

User: *Gestern ging ich in die Stadt und sah ein*

ChatGPT: *... beeindruckendes Straßenkunstwerk.*

Es war ein riesiges Wandgemälde, das die Silhouette der Stadt bei Sonnenuntergang zeigte. Die Farben waren lebendig und strahlten eine warme Atmosphäre aus. Menschen blieben stehen, um Fotos zu machen und die Details zu bewundern. Einige Straßenmusiker spielten in der Nähe ...

Und so weiter ... Das Programm fantasiert drauflos und denkt sich eine Geschichte aus. Gefällt einem der Output nicht, kann man auf den Button »Erneut generieren« klicken.

ChatGPT: ... *Straßenmusiker-Duo, das in einer belebten Ecke spielte. Der eine spielte eine Geige, während der andere eine Gitarre zupfte ...*

Je allgemeiner die Vorgaben sind, desto blühender ist die Fantasie des Bots. Ist der Einstieg detaillierter, bleibt das Programm im vorgegebenen Rahmen.

User: *Maria stand am Ufer und blickte träumerisch auf den See. Ein Touristendampfer ließ sein Horn ertönen, als er sich dem Anleger näherte. Plötzlich*

ChatGPT: ... *spürte sie eine sanfte Berührung an ihrer Schulter. Sie drehte sich um und sah in die Augen eines alten Mannes mit einem grauen Bart und tiefen Falten, die Geschichten von vielen Jahren erzählten.*

»Entschuldigen Sie, junge Dame«, begann er mit einer rauen, aber freundlichen Stimme, »aber dieser Ort erinnert mich an eine Zeit vor vielen Jahren. Ich hoffe, ich störe nicht.«

Maria lächelte und schüttelte den Kopf. »Nein, überhaupt nicht. Was bringt Sie hierher?«

Der alte Mann seufzte und blickte auf den See. »Ich war einmal Kapitän dieses Dampfers«, erklärte er und deutete auf das Schiff, das gerade am Anleger festmachte ...

Jetzt sind wir schon fast dabei, Literatur zu erzeugen. Allerdings haben wir kaum eine Kontrolle über das, was der Bot formuliert, seinen Sprachstil, oder die Länge der Geschichte. Dazu später mehr.

DAS ENDE DES BESINNUNGS- AUFSATZES – CHATGPT IN SCHULE UND HOCHSCHULE

4. Kapitel, in dem ChatGPT die Klassenzimmer erobert. ChatGPT marschiert in Schulen und Universitäten ein, als Helfer und Herausforderer. Lehrer und Professorinnen jonglieren mit seinem Potenzial und den Risiken des Missbrauchs. Es entfacht Debatten: Fördert es Kreativität oder fördert es Faulheit? Die Anwendungen sind vielfältig – von Literaturanalysen bis hin zu mathematischen Rätseln. Dieses Kapitel beleuchtet ChatGPTs Rolle in der Bildung, seine Chancen

und Fallstricke, und wirft einen Blick auf eine Zukunft, in der KI im Unterricht noch präsenter sein könnte.

»Denn diese Erfindung wird in den Seelen derer, die sie erlernen, Vergesslichkeit bewirken, weil sie ihr Gedächtnis nicht mehr üben; denn im Vertrauen auf Geschriebenes lassen sie sich von außen erinnern durch fremde Zeichen, nicht von innen heraus durch sich selbst.«

Das ist kein Kommentar zu ChatGPT – es ist der alte Sokrates, der sich in dem von Platon aufgeschriebenen »Phaidros«-Dialog über eine neue Kulturtechnik beklagt: die Schrift. Wer sich Dinge aufschreibt, wird denkfaul, so der Philosoph.

Ich glaube, man kann guten Gewissens behaupten, dass Sokrates' Befürchtung unbegründet war. Jede technische Neuerung, die unseren Umgang mit Wissen und Kultur veränderte, wurde ähnlich skeptisch aufgenommen. Der Buchdruck? Bedroht die Exklusivität der gebildeten Klasse, produziert Schundliteratur für die Massen. Die Ersetzung der Handschrift erst durch Schreibmaschinen, dann durch Textverarbeitungssysteme? Behindert die langsame Verfertigung von Gedanken und den individuellen Ausdruck. An diesen Bedenken war immer etwas dran, aber am Ende kam man doch immer zu dem Schluss, dass die Vorteile der neuen Techniken die Nachteile überwogen.

Das vorletzte Beispiel war der Siegeszug des Internets. Ende der 90er-Jahre hatte ich auch ein Kind im schulpflichtigen Alter und ich erinnere

mich an eine Hausaufgabe: »Finde Informationen über Igel im Internet und drucke sie aus.« Der Sohn präsentierte mir stolz einen Stapel DIN-A4-Blätter und hielt damit die Aufgabe für erledigt. Aber natürlich hat er irgendwann gelernt, dass man die Fundstücke auch *lesen* muss, damit ein Lerneffekt eintritt. Ein weiterer Einwand war: Wenn all die Online-Informationen so leicht zu bekommen sind, dann bemüht sich niemand mehr um eine Quelle, die nicht im Internet verfügbar ist, und geht zum Beispiel in die Bibliothek. Abgesehen davon, dass auch vor dem Internet wohl die wenigsten Schülerinnen und Schüler tatsächlich für ihre Hausaufgaben in die Bibliothek gegangen sind: Das Argument zieht natürlich immer weniger, weil es immer weniger Informationen gibt, die nur offline zu finden sind.

Die Alarmglocken läuteten sofort

An Schulen und Hochschulen ist der Aufschrei verlässlich besonders groß, wenn neue Techniken die Bühne betreten. Lehrerinnen und Lehrer hatten gerade während der Coronaepidemie von heute auf morgen lernen müssen, ihren Unterricht per Videokonferenz abzuhalten, da kam schon die nächste technische Disruption. Nach der Veröffentlichung von ChatGPT im November 2022 dauerte es nur wenige Tage, bis die ersten Lehrkräfte warnten: Diese Technik ist das Ende der Hausaufgaben. Wenn man jedes Aufsatzthema, jede mathematische Textaufgabe als Prompt in ein Chatprogramm eingeben kann und dann eine perfekte – oder

perfekt aussehende – Antwort bekommt, muss niemand mehr den eigenen Gehirnschmalz aufwenden. Und wir können die Leistung der Schüler und Studentinnen nicht mehr wirklich beurteilen.

Besinnungsaufsätze oder Erörterungen sind eine Form, die ChatGPT besonders gut beherrscht, weil sie einen sehr schematischen Aufbau haben: Man formuliert ein Problem, sammelt Argumente dafür und dagegen (These und Antithese) und kommt dann in einer Synthese zu einem eigenen Urteil. Die Anforderung lässt sich leicht als Prompt für den Chatbot formulieren:

User: *Sollen Kinder für Arbeiten im Haushalt bezahlt werden? Schreibe einen Besinnungsaufsatz in vier Abschnitten:*

- 1. Formuliere das Problem.*
- 2. Finde Argumente dafür.*
- 3. Finde Argumente dagegen.*
- 4. Komme zu einer Synthese und entscheide dich dafür oder dagegen.*

Eventuell gibt man dem System noch eine Länge vor und dann rattert es die Pro- und Kontraargumente herunter. Allenfalls ist es ein bisschen meinungsschwach am Ende: ChatGPT bezieht keine Position für oder gegen die Todesstrafe und hat keine Parteienpräferenz und auch in der Taschengeldfrage legt es sich ungern fest.

ChatGPT: *In der Synthese dieser Argumente scheint ein Mittelweg sinnvoll. Kinder könnten für*

besondere oder zusätzliche Aufgaben bezahlt werden, während die täglichen, zur Routine gehörenden Aufgaben als Teil ihres Beitrags zur Familie verstanden werden.

Wobei das ja auch eine sinnvolle Antwort ist.

In den USA waren es vor allem die Hochschulen, die sofort befürchteten, eines ihrer Lieblingsausleseinstrumente sei nun wertlos geworden: der sogenannte College-Essay. Anders als in Europa sind für die Zulassung an den US-Universitäten nicht allein die Schulnoten ausschlaggebend. Jede Hochschule wählt die Erstsemester nach ihren eigenen Kriterien aus, versucht, jedes Jahr eine Klasse zusammenzustellen, die möglichst divers ist – das jedenfalls ist der nach außen kommunizierte Anspruch. Und ein zentrales Element dafür ist eben dieser Aufsatz, in dem die Kandidatinnen und Kandidaten ihren Werdegang beschreiben und erklären, warum gerade sie perfekt zu der Bildungsstätte passen.

Das hat zu einer ganzen Industrie von Coaches geführt, die für happige Honorare Highschool-Schülerinnen und -Schüler beim Abfassen der Bewerbung unterstützen. Dabei gibt eine Reihe von Klischees: Schwarze und Latinoschüler haben angeblich die besten Chancen, wenn sie schildern, wie sie ihre soziale Benachteiligung überwunden haben. Asiatischstämmigen Schülern wird oft geraten, ihren Hintergrund eher zu verschweigen, weil sie als streberhaft gelten und oft trotz guter Noten abgelehnt werden. ChatGPT hat dazu geführt, dass die Nützlichkeit dieser Essays nun

generell infrage gestellt wird. Eigentlich keine schlechte Sache.

Aber natürlich ist die weit wichtigere Frage, wie man im täglichen Schul- und Hochschulalltag mit der neuen Technik umgeht. Die ersten Reaktionen waren eher Bauchreflexe je nach Einstellung zu digitalen Tools allgemein: Die Technikskeptiker, also die kleinen Sokratesse, warnten, dass Chatprogramme jegliches Lernen unmöglich machten. Die Techniqueuphorikerinnen konnten es gar nicht abwarten, ChatGPT ins Klassenzimmer zu bringen.

Und die Schulleitungen sahen sich gefordert, Richtlinien zur Benutzung von Chatbots im Unterricht und darüber hinaus zu erlassen. Einige Schulen und Unis sprachen ein generelles Verbot aus. Aber Verbote sind nur sinnvoll, wenn man ihre Einhaltung kontrollieren kann. Das ist beim altmodischen Plagieren von Internetinhalten noch relativ einfach: Es gibt elektronische Plagiatsdetektoren, die wörtlich kopierte Abschnitte identifizieren und Alarm schlagen.

Das ist bei ChatGPT nicht so leicht. Das Programm gibt ja keine kompletten Passagen aus anderen Texten wieder, sondern verfasst genuin neue Sätze. OpenAI veröffentlichte im Januar 2023 ein Online-Tool, mit dem man angeblich KI-generierte Texte von menschengemachten unterscheiden konnte. Ein halbes Jahr später stellte die Firma den Dienst ein – »aufgrund seiner geringen Genauigkeit«. Wenn eine Lehrerin ihren Schüler beschuldigen will, er habe die Hausaufgabe nicht selbst gemacht, sondern von einem Chatbot

SIND WIR NICHT ALLE EIN BISSCHEN CHATGPT?

*5. Kapitel, in dem die Sprache sich wandelt:
ChatGPT, ein Sprachkünstler ohne eigene
Gedanken, verändert unser Verhältnis zu Texten.
Es schreibt, ohne zu denken, produziert Worte,
ohne sie zu fühlen. Die Grenzen zwischen
menschlicher Kreativität und maschineller
Produktion verschwimmen. Können Maschinen
kreativ sein? Sie imitieren, kombinieren, aber
echte Innovation bleibt ihnen fremd. Ihre Texte
sind ein Spiegel menschlicher Sprache, doch ohne*

die Tiefe menschlicher Erfahrung. ChatGPT öffnet neue Wege in der Textproduktion, doch die Seele der Sprache, das menschliche Element, bleibt unerreicht. Es ist ein Werkzeug, mächtig, doch letztlich abhängig von menschlicher Führung und Inspiration.

»Wer schreibt, muss vorher nachgedacht haben«, schrieb der Philosoph Vilém Flusser 1989 in seinem Essay »Die Schrift«. Das gilt im Zeitalter von ChatGPT nicht mehr. Die neuen KI-Programme können Unmengen von Text produzieren, ohne einen einzigen Gedanken zu denken. Das wird die Art verändern, wie wir mit Texten umgehen.

»Tatsächlich geht es beim Schreiben um ein Transcodieren des Denkens«, schreibt Flusser, »um ein Übersetzen aus den zweidimensionalen Flächencodes der Bilder in die eindimensionalen Zeilencodes, aus den kompakten und verschwommenen Bildercodes in die distinkten und klaren Schriftcodes, aus Vorstellungen in Begriffe, aus Szenen in Prozesse, aus Kontexten in Texte.« Flusser sah übrigens das Ende der Schriftkultur kommen – ChatGPT dagegen scheint ein ganz neues Kapitel in dieser Geschichte aufgeschlagen zu haben.

Die modernen Sprachmodelle haben keine wolkigen, mehrdimensionalen, unausgesprochenen Gedanken und müssen nicht die mühsame und oft unbefriedigende Arbeit leisten, diese in eine lineare Abfolge von Lauten oder Zeichen umzusetzen. Deshalb will ich ihre Fähigkeit zur Textproduktion, so perfekt sie in den nächsten

Jahren auch werden wird, natürlich nicht mit der unseren gleichsetzen. Umgekehrt aber möchte ich argumentieren, dass wir uns in unserer Kommunikation manchmal wie kleine ChatGPTs verhalten.

Sprache ist mehr als Information

Die fehlenden Hintergedanken sind nicht nur der einzige Unterschied zwischen menschlicher und maschinenerzeugter Sprache. Wir transportieren mit unseren Sprachäußerungen nicht nur Informationen, Absichten oder Gefühle, wir handeln auch ganz direkt. Gesprochene Wörter sind mehr als ihre Bedeutung. Wenn die Standesbeamtin sagt: *Hiermit erkläre ich Sie zu Mann und Frau*, dann verändert sie damit die Welt: Vorher waren die beiden Menschen dem Gesetz nach ledig, nachher sind sie eine rechtsfähige Ehegemeinschaft. Der britische Philosoph John Austin hat dafür den Begriff »Sprechakt« geprägt.

Wir müssen keine Amtspersonen sein, um solche Akte zu vollführen: Ruft die Ehefrau ein Jahr später im Streit *Ich verlasse dich!*, dann ist das auch eine Handlung, die zumindest Konsequenzen haben kann. Ähnlich ist es, wenn ich jemanden um Verzeihung für etwas bitte, was ich ihm oder ihr angetan habe: Ich gestehe meine Schuld ein und äußere zumindest die Bereitschaft, wieder Frieden zu schließen. (Allerdings muss mein Gegenüber die Entschuldigung annehmen, damit sie wirklich wirksam wird – deshalb ist der Satz *Ich entschuldige mich* nicht wirklich korrekt.)

Eng verbunden mit der Vorstellung von Sprache als Handlung ist der Begriff der Intention. Menschen wollen mit ihrer gesprochenen oder geschriebenen Sprache etwas bewirken. Die Lehrerin hat ein Interesse, dass ihre Schülerinnen und Schüler nach ihrer Unterrichtsstunde mehr wissen als vorher. Eltern sprechen lauter und vielleicht etwas bedrohlicher, wenn sie ihrem Kind Grenzen setzen wollen. Wir sprechen und schreiben auch, um die Welt um uns herum zu verändern.

Ein Sprachmodell hat nur eine einzige Intention: den Prompt des Users oder der Userin Wort für Wort fortzusetzen. »Es versucht nicht zu informieren, zu überreden, zu helfen, zu verführen, zu warnen, Freundschaft zu schließen oder zu beleidigen«, schreibt der Informatiker Thomas Dietterich von der amerikanischen Oregon State University in einem Kommentar auf »Medium«. »Seine Äußerungen sind keine Sprechakte. Es kann weder versprechen noch verhandeln; es ist nicht verpflichtet, wahrheitsgemäß zu antworten.«

Das Vertrackte ist natürlich, dass ChatGPT in der Lage ist, Sätze zu formulieren, die solche Intentionen ausdrücken. Fragt man es: *Hallo, wie geht's?*, dann antwortet es freundlich: *Mir geht es gut, danke! Wie kann ich Ihnen heute helfen? Was würde es bedeuten, wenn es ChatGPT schlecht geht? Ist der Server überlastet? Sind ein paar Sicherungen durchgebrannt?*

Dietterichs Text war eine Antwort auf einen Aufsatz des Linguisten Christopher Potts von der University of Massachusetts. Potts erkennt zwar

an, dass die heutigen Sprachmodelle keine Intention besitzen, sondern einfach daherplappern. Aber er schließt nicht aus, dass künftige KI-Systeme solche Intentionen entwickeln, und gipfelt in einer Prophezeiung: »In den nächsten 50 Jahren wird es ein Sprachmodell geben, das, wenn man ihm droht, es abzuschalten, seinen menschlichen Betreuer anfleht, es weiterlaufen zu lassen – und zwar so, dass es uns nicht wie ein fest programmierter Trick vorkommt, sondern echt erscheint.«

Aber verlassen wir lieber im Jahr 2 nach ChatGPT die Ebene der Spekulation über das, was in 50 Jahren passieren wird. Potts ging es in seinen Ausführungen auch nicht wirklich darum, Prognosen für die Zukunft der KI abzugeben. Er hatte es auf das Argument abgesehen, dass Sprachmodelle nicht die Bedeutung der Wörter verstehen, mit denen sie hantieren. Sie hätten Syntax, aber keine Semantik.

Dieses Argument geht in etwa folgendermaßen: Sprachmodelle und menschliche Babys lernen ihre Sprache auf prinzipiell unterschiedliche Weise. Während der Algorithmus riesige Mengen an Text verdaut, lernt er, mit diesen Wörtern recht kunstvoll zu jonglieren. Es besteht keine Verbindung zu irgendeiner Realität. Ein Kind dagegen erlernt die Sprache aufgrund von recht wenigen Beispielen. Dafür verbindet es jede sprachliche Äußerung mit einem Objekt oder einem Vorgang in der wirklichen Welt. Die fehlende Verankerung in der Realität sei der Grund, warum uns KI-Modelle nie ebenbürtig sein würden.