

Informationen zu ChatGPT und vergleichbaren Werkzeugen

Karsten Wolf

19.7.2023

ChatGPT, deepL und andere Programme sind KI-basierte Werkzeuge zur Textgenerierung. Hier sind die wichtigsten Informationen am Beispiel von chatGPT kurz zusammengefasst.

Wie funktioniert chatGPT?

Das Programm generiert einen Text, indem es Wort für Wort (zum Teil sogar nur Wortteile) aneinanderreihet. Beginnend mit der Anfrage selbst, wird abhängig vom bereits festgelegten Textteil für jedes Wort eine Wahrscheinlichkeit dafür ermittelt, dass dieses Wort in der Antwort als nächstes auftritt. Das Wort mit der höchsten Wahrscheinlichkeit wird an den Text angefügt. Dieser Prozess wird solange wiederholt, bis einem speziellen Symbol „Beende den Text“ die höchste Wahrscheinlichkeit zugewiesen wird. Einige Werkzeuge können mehrere Antwortvarianten zur gleichen Anfrage erzeugen, indem sie auch Wörter mit etwas geringeren Wahrscheinlichkeiten in den Text einbauen.

Die Wahrscheinlichkeiten wurden anhand von umfangreichen Trainingsdaten ermittelt und durch menschliche Bewertungen zahlreicher Antworten optimiert. Wesentlichen Einfluss auf die Wahrscheinlichkeiten hat weiterhin ein fortschrittliches Sprachmodell, in dem nicht nur grammatische, sondern auch semantische und stilistische Beziehungen abgebildet sind. So kann unterschieden werden, ob es sich bei einem „König“ um einen Herrscher oder eine Schachfigur handelt. Die Unterscheidung wird durch Begriffe im Kontext (z.B. „Rochade“, „Läufer“, „Prinzessin“, „Schloss“) ermöglicht. Für unterschiedliche Könige gibt es dann auch unterschiedliche Wahrscheinlichkeiten zur Fortsetzung des Textes („den König ziehen“ versus „den König ersuchen“).

Woher stammen die Trainingsdaten?

Es sind nicht alle Quellen bekannt. Zu den Quellen gehören aber auf jeden Fall zahlreiche frei verfügbare Textdatenbanken wie z.B. das Gutenberg-Projekt. Bei chatGPT wurde das Training vor einiger Zeit beendet. Damit ist chatGPT kaum in der Lage, sinnvoll auf Anfragen zu aktuellen Themen zu reagieren. Andere KI-Werkzeuge lernen während der Benutzung weiter. Dieses Weiterlernen birgt Risiken. So musste Microsoft vor einiger Zeit einen KI-basierten Chatbot schon 24 Stunden nach Start vom Netz nehmen, da dieser durch die Nutzer darauf trainiert worden war, sich vornehmlich rassistisch und rechtsradikal zu äußern.

Was ist der Daten-Bias?

Die Trainingsdaten sind thematisch nicht ausgewogen. Auch bei den bewertenden Menschen sind nicht alle Gruppen gleichmäßig repräsentiert. So gibt es z.B. einen Überhang bei weißen, männlichen US-Amerikanern im Alter zwischen 25 und 45 Jahren. Den Effekt kann man gut beobachten, wenn man z.B. auf eine Anfrage zu den „Rostock Dolphins“ eine Antwort erhält,

die von American Football statt, wie es korrekt wäre, von Handball handelt. Ursache ist die Überrepräsentation von Daten zum Football-Verein „Miami Dolphins“.

Sind Antworten von chatGPT vertrauenswürdig?

Um Himmels Willen, nein! Dies ist nicht einmal der Fall, wenn wir die bereits beschriebenen Probleme (Horizont der Trainingsdaten, Daten-Bias) herausrechnen. Wie bei jedem KI-Werkzeug unterliegen lediglich die Trainingsdaten einer wie auch immer gearteten Qualitätssicherung (zum Vergleich: bei Wikipedia sind *alle* Texte qualitätsgesichert, wenn auch auf schwankendem Niveau). Antworten müssen aber nicht einmal bei den Trainingsdaten mit den Originalantworten übereinstimmen, da die Wahrscheinlichkeiten für die Gesamtheit der Trainingsdaten optimiert sind. Für sämtliche Anfragen, die nicht exakt mit Trainingsdaten übereinstimmen, gilt das Prinzip Hoffnung: ähnliche Fragen haben doch bestimmt ähnliche Antworten, oder? Die Chance, dass diese Hoffnung bestätigt wird, ist umso größer, je engmaschiger zu einem bestimmten Thema Trainingsdaten vorliegen. Deshalb besteht chatGPT einige stark standardisierte Tests wie z.B. College Entrance Exams. ChatGPT eignet sich sicher sehr gut als Formulierungsvorschlag. Die Texte erfordern aber ein sehr, sehr gründliches inhaltliches Lektorat.

Müssen wir rechtliche Rahmenbedingungen anpassen?

Nein. Werkzeuge wie chatGPT sind, juristisch betrachtet, *Hilfsmittel*, also gleichgestellt zu bekannten Hilfsmitteln wie z.B. Taschenrechnern. Daher kann chatGPT im vorhandenen rechtlichen Rahmen adressiert werden. Solange sich Lehrende nicht explizit zu chatGPT äußern, ist die Nutzung nicht erlaubt.

Dennoch gibt es sinnvolle Maßnahmen zur rechtlichen Flankierung.

Eine dieser Maßnahmen besteht darin, alternative Prüfungsformen zu ermöglichen. Hierzu steht das Rektorat im Kontakt zum WKM mit dem Ziel, die Prüfungsform „Referat im Seminar“ vom nicht handhabbaren Erfordernis eines zweiten Prüfers (festgelegt im Landeshochschulgesetz) zu befreien.

Wenn Verantwortliche wegen chatGPT zu alternativen Prüfungsformen wechseln möchten, sind Satzungsänderungen in den SPSO nötig.

Das WKM ist in einem Fachgespräch mit Expertinnen und Experten zum Thema im März 2023 zu der Auffassung gelangt, dass auch landesrechtliche Regelungen (z.B. LHG) keiner unmittelbaren Anpassung wegen chatGPT bedürfen.

Die EU plant Regulierungen und einen Code of Conduct für den Einsatz KI-basierter Werkzeuge.

Kann ich die Nutzung von chatGPT verbieten?

Juristisch ist dies möglich. Es obliegt der zuständigen Lehrperson, die erlaubten (und somit auch die verbotenen) Hilfsmittel für die Anfertigung einer Prüfungsleistung festzulegen. Die Nutzung von chatGPT kann aber oft nicht mit der für prüfungsrechtliche Sanktionen erforderlichen Sicherheit nachgewiesen werden. Daher ist von einem Verbot eher abzuraten.

Ein wirksames Verbot ist bei Prüfungsleistungen in Präsenz möglich. Sollte die zu prüfende Kompetenz mit der Nutzung von KI-Werkzeugen unverträglich sein, ist ein Wechsel auf komplett andere Prüfungsformen (vorzugsweise in Präsenz) zu erwägen.

[Kann ich ohne Moduländerung die Nutzung von chatGPT erlauben?](#)

Ja. Die für eine Prüfungsleistung erlaubten Hilfsmittel sind durch die Lehrperson rechtzeitig festzulegen. Zwar sind alle nicht explizit erlaubten Hilfsmittel automatisch verboten, jedoch empfiehlt es sich in der aktuellen Situation, klare Aussagen zu KI-Werkzeugen zu treffen, sie also explizit zu erlauben oder zu verbieten.

[Gibt es Vorgaben des Rektorats?](#)

Nein. Die Entscheidung darüber, ob und in welcher Form Sie chatGPT in Ihren Lehrveranstaltungen bzw. Prüfungen nutzen, obliegt Ihnen im Rahmen Ihrer Freiheit der Lehre. Nur Sie können entscheiden, inwieweit Ihre Kompetenzziele mit oder ohne chatGPT erreichbar sind. Wollen Sie in Lehrveranstaltungen chatGPT durch Studierende nutzen lassen, ist allerdings dem Datenschutz Rechnung zu tragen (siehe Frage weiter unten).

Wollen Sie wegen chatGPT die Prüfungsform wechseln, ist eine Satzungsänderung nötig, die durch Ihre Fakultät auf den Weg zu bringen ist. Ausgestaltungen innerhalb der aktuellen Prüfungsform (chatGPT erlauben, Nutzungsprotokoll verlangen) können Sie in eigener Regie vornehmen, müssen diese lediglich rechtzeitig (in der Regel zu Semesterbeginn) bekanntgeben.

Das Rektorat steht Ihnen bei Ihren Fragen zum Einsatz von chatGPT gern beratend zur Seite.

[Wie kann ich chatGPT sinnvoll beim Schreiben meiner Arbeit verwenden?](#)

Zunächst: Wenn die Nutzung nicht erlaubt ist, lass es! Erstens wäre es illegal. Zweitens: Wenn Du die Angebote im Studium, eine Fertigkeit zu trainieren, nicht annimmst, dann fehlt Dir dieses Training später und Du bist schlechter auf Deinen Beruf vorbereitet. Wer beim Lauftraining schummelt, ist auch im Wettkampf langsam.

Ist die Nutzung erlaubt, kann chatGPT Dir zunächst beim Lösen bestimmter Schreibblockaden (z.B. der „Angst vor dem weißen Blatt“) helfen. Zu Deiner Anfrage generiert es einen ersten Text, den Du dann redigieren, ausbauen und weiterverarbeiten kannst. Außerdem können Dir Antworten Anregungen, Inspirationen, neue Blickwinkel auf Dein Thema liefern, so dass Du in neue Richtungen recherchieren und arbeiten kannst. Vielleicht findest Du auch eine gute Formulierung eines Gedankens. Wichtig ist, dass Du alle Behauptungen mit seriöser Quellenarbeit nachprüfst. Der Text, den chatGPT Dir liefert, unterlag keiner Qualitätssicherung, keiner Plausibilitätsprüfung, keinem Faktencheck. Am Ende musst Du für Deinen Text geradestehen.

[Wie kann ich die Leistung meiner Studierenden bewerten, wenn doch die KI den Text geschrieben hat?](#)

Auch die Nutzung von KI-Werkzeugen erfordert intellektuelle Leistungen:

- Inhaltliches und stilistisches Lektorat
- Geeignete Formulierung und schrittweise Verfeinerung von Anfragen
- Weiterverarbeitung des generierten Textes
- Sicherstellung der Gesamtstruktur

Eine komplette schriftliche Arbeit erfordert normalerweise eine komplexe Interaktion mit der Software sowie eine tiefgründige Auseinandersetzung mit den gelieferten Antworten. Die Schöpfungshöhe können Sie zum Beispiel dadurch einschätzen, dass Sie als Anhang zur gelieferten Arbeit die Abgabe des vollständigen Nutzungsprotokolls mit dem KI-Werkzeug verlangen. Eine weitere Möglichkeit besteht darin, die Hausarbeit um ein Kolloquium zu ergänzen, um die tatsächliche Auseinandersetzung mit dem gestellten Thema zu überprüfen.

Wie erkenne ich die Nutzung von chatGPT in einer abgegebenen Arbeit?

Eine sichere Erkennung ist nicht möglich. Vor allem ist damit zu rechnen, dass bekannte Schwächen aktueller Versionen in absehbarer Zeit behoben werden. Zu den aktuell noch erkennbaren Schwächen gehören:

- Generierte Texte behaupten durchaus häufig kompletten Unsinn oder driften vom eigentlichen Thema ab
- Der Daten-Bias und der Trainingshorizont sind gelegentlich erkennbar
- Zitate werden oft inkorrekt wiedergegeben
- Referenzen sind häufig reine Fantasie

Unsinn werden Sie umso häufiger sehen, je weiter das Thema vom Mainstream der Trainingsdaten entfernt ist. Die Fehler bei Zitaten und Referenzen sind durch das Grundprinzip des Wort-für-Wort-Vorgehens verursacht. Künftige Entwicklungen werden Zitate und Referenzen vielleicht als untrennbare Bausteine verwalten und entsprechende Probleme beheben.

Bei einer komplett unlektorierten Textübernahme werden Sie die Nutzung von chatGPT vermutlich derzeit noch anhand der genannten Schwächen erkennen können. Sie werden auch plötzliche „Sprünge“ bei den Formulierungsfertigkeiten einiger Ihrer Studierenden feststellen. Mit etwas Aufwand lassen sich aber die Texte so redigieren, dass Sie die Nutzung der KI kaum noch nachweisen können.

Die Bezahlversion von chatGPT ist deutlich leistungsfähiger als die freie Version. Ihr liegen umfangreichere Trainingsdaten und weiterentwickelte Algorithmen zugrunde. Entsprechend schwerer wird es sein, generierte Texte als solche zu erkennen.

Was muss ich bei der Korrektur einer Hausarbeit beachten?

Falls eloquente Formulierungen je ein Indiz für Sachkenntnis waren, sind diese Zeiten nun vorbei. Sie sollten die einzelnen Fakten, jedes Zitat, jede Referenz genau prüfen. Die Schöpfungshöhe können Sie am besten anhand des Protokolls der Nutzung des Werkzeuges ermitteln.

Wie kann ich geeignete Hausarbeitsthemen stellen?

Vorteilhaft sind Themenstellungen, zu denen eine geringe Dichte an Trainingsdaten vorliegt. Das sind Themen mit aktuellen Bezügen, abseits vom Mainstream, die ungewöhnliche Aspekte betreffen. Ist dies im Fachkontext nicht möglich, sollte die Prüfungsform „Hausarbeit“ grundsätzlich überdacht werden.

Kann ich Texterstellung mit chatGPT in meiner Lehrveranstaltung einüben?

Es spricht nichts dagegen. Wir entlassen unseren Absolventinnen und Absolventen in eine Welt, in der chatGPT und andere Werkzeuge nun einmal existieren. Das Wissen um sinnvolle Einsatzszenarien und die notwendige Kritikfähigkeit im Umgang mit den Werkzeugen sind vermutlich hilfreich.

Erstellt chatGPT Plagiate?

Nein, normalerweise werden keine Plagiate generiert. Dies hängt mit der Wort-für-Wort-Strategie beim Generieren von Texten zusammen. Mit jedem Wort steigt die Wahrscheinlichkeit, von einem Text aus den Trainingsdaten abzuweichen. Daher ist auch Software für Plagiaterkennung nicht in der Lage, mit chatGPT erstellte Texte zu überführen.

Verletzt chatGPT Urheberrechte?

Es ist mir nicht bekannt, ob beim Training urheberrechtlich geschützte Daten verwendet wurden. Selbst dann ist juristisch umstritten, ob eine Verletzung des Urheberrechts vorliegt. Das Urheberrecht gestattet die vollständige statistische Auswertung eines Textes (z.B. das Zählen von Häufigkeiten von Wörtern). Es ist noch nicht geklärt, ob es sich beim Training der KI um eine solche statistische Auswertung handelt. Gerichtliche Auseinandersetzungen zu diesem Thema scheinen in Vorbereitung zu sein.

Wer ist Urheberin bzw. Urheber eines KI-generierten Textes?

Klar ist: die Software ist es nicht. Klar ist auch: die Herstellerfirma ist es nicht. Die Nutzerin bzw. der Nutzer eines KI-Werkzeuges kommt als Urheberin bzw. Urheber grundsätzlich in Frage. Dazu muss im Einzelfall geklärt werden, ob, z.B. durch die geschickte Formulierung der Anfragen, eine ausreichende Schöpfungshöhe zustande kam.

Wie steht es mit dem Datenschutz?

Schon bei der freien Version melden Sie sich mit E-Mail und einer Telefonnummer an (Zwei-Faktor-Authentifizierung). Ihre Anfragen und Verbindungsdaten gehen an die Betreiberfirma. Die Daten gehen zwar nicht in das (abgeschlossene) Training der aktuellen Version ein, werden aber möglicherweise für das Training zukünftiger Versionen eingesetzt. Stellen Sie daher niemals Anfragen, die schutzwürdige Daten beinhalten!

Können wir die Bezahlversion beschaffen?

Bedenken in Sachen Datenschutz bestehen. Eine Beschaffung zu Forschungszwecken hat bei einer klaren Beschreibung des Forschungsziels sowie von Regelungen zu Freiwilligkeit der Nutzenden und zum Verbot schutzwürdiger Daten in den Anfragen dennoch Aussicht auf Bewilligung.

Nächstes Ziel ist es, eine tragfähige Regelung zur Beschaffung durch bzw. für Lehrende zu finden. Lehrende sollen die Möglichkeit erhalten, sich mit den Möglichkeiten des Werkzeuges vertraut zu machen und Schlussfolgerungen in Bezug auf Themenstellungen und Bewertungen ziehen zu können.

Eine Bereitstellung der Bezahlversion für Studierende ist wegen der Datenschutzbedenken in näherer Zukunft nicht vorgesehen. Hier sind höhere Anforderungen zu stellen, für deren Erfüllung noch Ideen fehlen.

Neben dem Datenschutz bestehen ggf. zusätzliche Hindernisse für eine Beschaffung. Die Bewertung der riesigen Mengen an Trainingsdaten wurden bei einigen Anbietern in Länder mit niedrigem Gehaltsniveau ausgelagert, so dass sie die für die Beschaffung notwendige Erklärung zum Mindestlohn nicht abgeben können.

Wo finde ich weitere Informationen?

Webseiten zu chatGPT:

<https://openai.com/blog/chatgpt>

<https://chatopenai.de>

Rechtliche und didaktische Hinweise, erarbeitet an der Ruhr-Uni Bochum:

<https://hss-opus.ub.ruhr-uni-bochum.de/opus4/frontdoor/index/index/docId/9734>

Eine umfangreiche Link-Sammlung zum Thema:

https://www.mmkh.de/fileadmin/veranstaltungen/netzwerk_landesinitiativen/KI-Generatoren/2023-03-14_KI-Generatoren_Hochschullehre_Links.pdf