

„Der neue Gott ist nackt!“

ChatGPT im Bildungswesen

| CLEMENS H. CAP | Seit der Freigabe der Technologie ChatGPT wird diskutiert, welche Konsequenzen sie für das Lernen in Schulen und Hochschulen hat. Drei Anwendungsbeispiele, die grundsätzliche Fragen aufwerfen.

Nach einer vielzitierten Studie benötigte Instagram zweieinhalb Jahre und TikTok neun Monate, um die Marke von 100 Millionen aktiven Nutzern pro Monat zu erreichen, bei ChatGPT waren es nach der ersten Freigabe am 30. November 2022 weniger als drei Monate. Der Hype um ChatGPT ist kaum zu schlagen. Es kann programmieren, kennt den zweiten Hauptsatz der Thermodynamik und eignet sich für die Urlaubsplanung. Die Technologie kann sogar für Ehekrisen sorgen: ChatGPT erklärte neulich einem Reporter der *New York Times* seine Liebe und forderte ihn auf, sich von seiner Frau zu trennen. In der Wissenschaft soll die Recherche eine der wesentlichen Anwendungen des Systems sein, so liest man. ChatGPT besteht die Fragen eines medizinischen Staatsexamens in den USA sowie den Mathe- und Geschichtsteil des bayrischen Abiturs, in Minnesota und Pennsylvania besteht es juristische Prüfungen und erhält in Wharton sogar ein B- bei einer MBA Prüfung. Nun schnellen auf Google die Suchanfragen zu ChatGPT hoch. Ministerien in England erwägen den Einsatz in der Regierungsarbeit,

das Bundesforschungsministerium verwendet das System bereits bei parlamentarischen Anfragen. Der deutsche Ethikrat äußert sich euphorisch über die Nutzung von ChatGPT in der Medizin. Ich lese mit, bewundere und stimme in die Anbetung des neuen Gottes mit ein. Eigene erste Gehversuche überzeugen meine Frau, eine Englischlehrerin: ChatGPT erklärt die Anwendung des *gerund* korrekt. Auch die wesentlich anspruchsvollere Frage nach den *stylistic devices* in Oscar Wildes Novelle „The Picture of Dorian Gray“ beeindruckt nachhaltig: ChatGPT nennt Fachbegriffe, gibt konkrete Beispiele aus dem Text und nennt auf Nachfrage Seitenzahlen. Allein die Frage, nach welcher Edition hier zitiert wird, bleibt offen, das System verweist auf Werk Ausgaben im public domain.

Eigene Beobachtungen

Ich will es genauer wissen! Neugierig „wurstle“ ich an zwei Wochenenden etwas herum und komme völlig ernüchtert zum Schluss: Der neue Gott ist nackt!

Beispiel 1: Rechenaufgabe: Die Textaufgabe „ 1234×2345 “ beantwortet das System falsch. Frage ich auf Englisch, so erhalte ich eine andere, ebenso falsche Antwort. Bitte ich um den detaillierten Rechenweg, so nutzt das System dafür die klassische Methode der händischen Multiplikation mit korrekter Einrückung nach Stellenschreibweise und ergänzt die Tabelle mit sinnvollen Hilfestellungen, aber alle vier Einzelberechnungen sind falsch. Zudem bekomme ich jetzt eine dritte falsche Antwort.

Frage ich auf Englisch nach dem Rechenweg, so erstellt das System ein Schema für 1234×1234 , eine Aufgabe, die niemand gestellt hatte. Der erste Summand im falschen Schema ist allerdings richtig berechnet ($6170 = 5 \times 1234$). Immerhin: Ich erhalte ein bereits von früher bekanntes falsches Ergebnis und nicht noch ein viertes falsches Ergebnis. Jetzt weise ich das System auf seinen Fehler hin. ChatGPT entschuldigt sich freundlich und bedankt sich für meine Korrektur, erkennt allerdings nicht, dass ich ihm dabei böswillig selber ein falsches Resultat untergemogelt habe.

Beispiel 2: Recherche: Vielleicht sind Rechenaufgaben für einen Computer ja die falsche Anwendungsdomäne! Ich bitte das System nach meinem Vater Ferdinand Franz Cap zu recherchieren. Das System *erfindet* einen mir nicht nachweisbaren österreichischen Philosophen und Theologen, den in der Wikipedia seit 2018 verzeichneten Physiker findet es nicht. Auf die Bitte, dazu doch Quellen anzugeben, *erfindet* das System Referenzen auf zwei Originalarbeiten und drei Biographien. Da ich diese Fantasie-Werke nicht finde, bitte ich um mehr Details.

ChatGPT nennt Verlagshäuser und vier ISBN Nummern. Ich prüfe nach: Alle vier Nummern sind *frei erfunden*. Ich mache das System darauf aufmerksam. ChatGPT entschuldigt sich freundlich und liefert teilweise neue Nummern, aber wiederum keine einzige gültige ISBN. Ich protestiere. Nach einer freundlichen Entschuldigung, möglicherweise wären ja manche Nummern veraltet, liefert das System nun neuere ISBN-13 Formate. Eine davon verweist auf ein existentes Buch, jedoch mit gänzlich anderem als dem behaupteten Titel. Die Bitte, mir die Herkunft dieser

AUTOR



Clemens H. Cap ist Professor für Informations- und Kommunikationsdienste an der Universität Rostock.

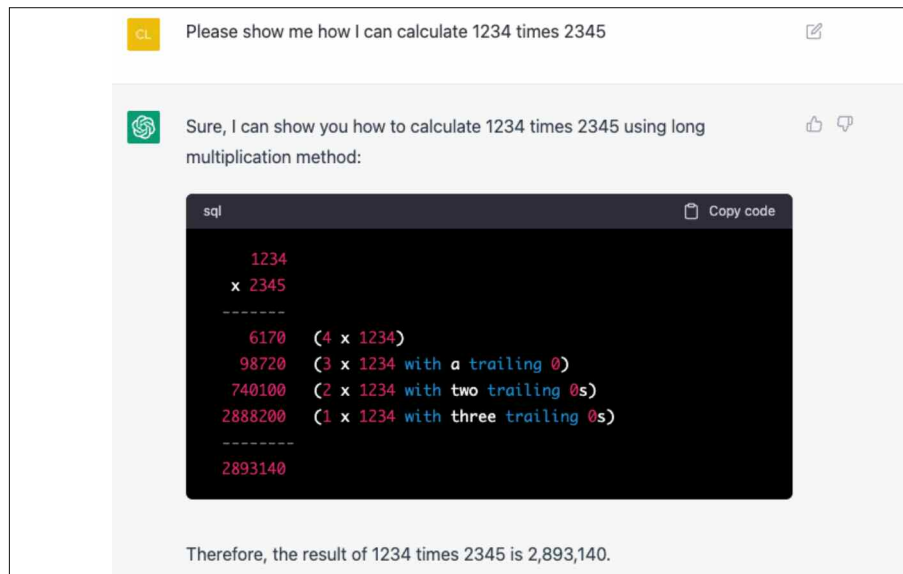


Abbildung: 1234×2345 berechnet nach ChatGPT.

Informationen zu nennen, führt zu einer Reproduktion derselben fiktiven Referenzen. Als Verwalter des wissenschaftlichen Nachlasses meines Vaters kenne ich sein Schrifttum. Die Erklärung des Versagens von ChatGPT fällt mir daher leicht.

Beispiel 3: Texterstellung: Ich bitte ChatGPT auf Deutsch um einen Aufsatz zu Quicksort, einem Informatik-Thema, geschrieben im Stil von Donald Trump, anschließend von John F. Kennedy. Das System antwortet auf Englisch, technische Details sind korrekt. Wer Reden dieser Präsidenten kennt, kann den Stil erkennen. Die Sache macht langsam Spaß! Ich bitte um einen Aufsatz im Stil von Angela Merkel. Das System wechselt die Sprache, wieder passt der Stil! Nun bitte ich um einen Aufsatz über Quicksort im Stil von Hankelbert Miller. Das System antwortet auf Englisch, im Stil eines Cowboys, bei großzügiger Bewertung ist es technisch korrekt. Sicher fragen Sie sich jetzt, wer dieser Herr ist. Ich weiß es nicht, denn ich habe den Namen frei erfunden. Da ChatGPT meine Frage zuerst aber beantwortet hat, frage ich das System. Die Phrasen des Bedauerns und der Entschuldigung kenne ich bereits. Das System gibt zu, die Person *frei erfunden* zu haben. Weitere acht Fragen, in denen ich das zunehmend kritisch hinterfrage, liefern wachsweiße, höfliche aber letztlich unbrauchbare Begründungs- und Betroffenheitslyrik. Den Satz von der „Anwendung von ChatGPT in Gutachten und Forschungsanträgen“ hatte ich im Kollegenkreis als Witz bereits mehrfach gehört!

Selbstbetrug

Die spannende Frage ist nicht, was ChatGPT kann, was nicht und warum. Die Themen, die uns beschäftigen sollten, sind soziologischer Natur.

Wie ist diese Begeisterung für ChatGPT in der Öffentlichkeit möglich? Mit welcher Motivation wird ein System freigegeben, das zwar offenkundig Multiplikationsaufgaben in Texten korrekt *erkennt*, dann aber vierstellige Zahlen *falsch multipliziert*? Wie kann ein System als „tauglich für Recherchen“ angesehen werden (sic!), das im Brustton der Überzeugung Zitate und Personen *völlig frei erfindet*? Woher kommen Vorschläge, solche Systeme auch in der Regierungsarbeit zu nutzen? Vor allem aber frage ich mich: *Wie zum Teufel* (sic!) kann ein solches Programm in unseren Bildungssystemen so erfolgreich punkten wie in der Einleitung beschrieben? Wo liegt hier das *echte* Problem? Wissenschaft, so lernte ich vor 35 Jahren, wäre ein hartes und trockenes Geschäft, das in sehr engen Bahnen verlaufe. Sie benötige ein hohes Maß an Selbstkritik und Disziplin, könne keine Wahrheiten finden und höchstens wackelige Thesen aufstellen, von denen sie die allermeisten im Laufe der Zeit wieder falsifiziere; alle ihre Prozesse benötigten extrem viel Zeit.

Wissenschaft, so lese ich seit 20 Jahren, wäre spannend und lustig und solle so kommuniziert und gelehrt werden. Digitalisierung, so lese ich seit zehn Jahren, verspräche die Lösung vieler Probleme. Künstliche Intelligenz schließlich, allen voran ChatGPT, so lese ich seit einem Monat, wäre ein revo-

lutionärer Schritt in die Zukunft, ein guter Kommunikator, vielleicht auch Lehrer oder gar Arzt.

Mit Wahrsagerei kenne ich mich nicht aus und den Schritt von Beobachtungen zu Theorien kenne ich als gefährlich aber interessant. Daher zum Schluss nur eine Beobachtung: Ich kann noch 1234×2345 von Hand ausrechnen – ChatGPT kann es nicht, liegt damit aber im Trend. In meiner letzten Klausur scheiterte rund ein Viertel der Erstsemester an der Nebenrechnung 36×36 . ChatGPT kann diese einfache Aufgabe jedoch. Die Schlagzeile für die Wissenschaftskommunikation, „ChatGPT bei Klausur besser als Erstsemester in Informatik“ ist von der Faktenlage her begründbar. In Kürze erhält das interne Modell von ChatGPT ein *update*. Vielleicht kann es dann auch 1234×2345 ausrechnen – selbst wenn (Vorsicht: Ironie) so komplizierte Aufgaben natürliche und künstliche Intelligenzen eigentlich überfordern und in Technologie und Medizin von heute bei falschen Ergebnissen „nur“ einstürzende Brücken oder Todesfälle verursachen können.

Völlig unklar aber ist mir die Debatte über Nutzung oder Verbot von ChatGPT im Bildungswesen. Wer ins Fitnesscenter geht zum Trainieren, strampelt am Fahrrad selber und schließt nicht einen Motor an, der das für ihn übernimmt. Wenn die gleiche Person dann nach Paris reist, muss sie das nicht unbedingt mit eigener Muskelkraft tun. Ebenso sollte es das Ziel jedes Trainers sein, die Athleten zu Höchstleistungen anzuspornen, und nicht primär Dopingkontrollen durchzuführen. Das Bild vom Selbstbetrug im Fitnesscenter lässt sich auf ChatGPT und das Bildungswesen übertragen – und das Argument gilt auch ganz unabhängig vom IQ künstlicher Intelligenzen und ihrer jeweiligen Updates und Konkurrenzprodukte.

Eine Gesellschaft, die nicht interessiert ist, selber auch durchschnittliche repetitive geistige Arbeit zu erlernen und zu trainieren, wird nie in der Lage sein, geistige Höchstleistungen zu vollbringen und sinnvolle Werkzeugnutzung stringent zu überwachen. Wenn Unterricht ausschließlich Spaß machen und interessant sein soll und nicht *primär* das Ziel hat, sinnvolle Leistungsfähigkeit zu vermitteln, so wird – ganz werbungsfrei – diese Leistungsfähigkeit schließlich verloren gehen.