

Können wir künstliche Intelligenzen wie ChatGPT produktiv für die medizinbibliothekarische Arbeit nutzen?

Can we use artificial intelligence like ChatGPT productively for medical library work?

Abstract

Since the American company OpenAI released the pre-release version of its chatbot ChatGPT in November 2022, a lively discussion has arisen about the benefits and use of speech AI in our society. Since its release ChatGPT has already fundamentally changed the understanding of written texts and it will continue to do so. This has sparked a debate about the use and impact of language AI in higher education. The potential applications of ChatGPT in higher education range from assistance with teaching and learning, to automation of routine tasks, to conversion and programming work. In the paper, examples from the medical library sector will demonstrate the power but also the limitations of ChatGPT.

Keywords: ChatGPT, linguistic AI, artificial intelligence, AI, chatbot, large language models (LLMs)

Zusammenfassung

Seit die amerikanische Firma OpenAI im November 2022 die Vorversion ihres Chat-Bots ChatGPT, veröffentlicht hat, ist eine lebhafteste Diskussion über die Vorteile und den Einsatz von Sprach-KI in unserer Gesellschaft entstanden. ChatGPT hat das Verständnis von geschriebenen Texten seit seiner Einführung bereits grundlegend verändert und wird es weiter verändern. Dies hat auch eine Debatte über den Einsatz und die Auswirkungen von Sprach-KI in der höheren Bildung ausgelöst. Die Anwendungsmöglichkeiten von ChatGPT im Hochschulbereich reichen von Unterstützung beim Lehren und Lernen über die Automatisierung von Routineaufgaben bis hin zu Konvertierungs- und Programmierarbeiten. In dem Beitrag sollen Beispiele aus dem medizinbibliothekarischen Bereich die Leistungsfähigkeit, aber auch die bis jetzt bestehenden Grenzen von ChatGPT aufzeigen.

Schlüsselwörter: ChatGPT, Sprach-KI, künstliche Intelligenz, KI, Chatbot, Sprachmodell

Einleitung

Im November 2022 hat das amerikanische Unternehmen OpenAI die Pre-Version des Chatbots ChatGPT veröffentlicht, der auf KI (künstlicher Intelligenz), dem Sprachmodell GPT-3-5 von OpenAI, beruht. Seitdem hat sich eine lebhafteste Diskussion über Nutzen von und Umgang mit Sprach-KI in unserer Gesellschaft entwickelt. ChatGPT hat mit dem ersten Tag das Verständnis von geschriebenen Texten verändert. Bereits nach wenigen Tagen hatten sich mehrere Millionen Nutzer registriert, was Social-Media-Plattformen wie Facebook, Instagram oder Twitter erst nach Monaten erreicht haben [1]. Man kann ohne

Übertreibung behaupten, dass die zum Teil kostenlose App wie eine Bombe eingeschlagen hat.

Was für die breite Öffentlichkeit wie eine sensationelle Neuheit aus dem Nichts kam, hat die deutsche Hochschullandschaft nicht gänzlich unvorbereitet getroffen [2]. An deutschen Hochschulen existieren zahlreiche KI-Studiengänge, die sich schwerpunktmäßig mit Anwendungen beschäftigen, bei denen Maschinen menschenähnliche Intelligenzleistungen erbringen [3]. Zudem beschäftigen sich viele Hochschullehrende seit Jahren mit dem Einsatz von KI im Kontext Hochschulbildung und Bildungsauftrag [2].

Stefanus Schweizer¹

Christian Vogel²

Oliver Weiner³

1 UB Mainz, Bereichsbibliothek
Universitätsmedizin

2 Medizinische Bibliothek der
Vinzenz Gruppe, Linz,
Österreich

3 UB Kiel, Zweigbibliothek
Medizin, Kiel, Deutschland

Die kritischen Stimmen, die von Datenschutzbedenken, Diskriminierungsgefahren, ethischen & urheberrechtlichen Fragen bis hin zu negativen Einflüssen auf den Arbeitsmarkt mit der Bedrohung von Arbeitsplätzen reichen, sollten nicht ignoriert werden. Eine Verteufelung oder ein striktes Verbot in der Hochschullehre erscheinen jedoch nicht sinnvoll, da die Sprach- und Textausgaben zum einen nicht immer eindeutig zu identifizieren sind und sich zum anderen nützliche Einsatzgebiete zeigen. Restriktionen im Hochschuleinsatz sind bei einer rasant fortschreitenden Technik nur schwer umsetzbar. In naher Zukunft werden Komponenten von Sprach-KI in Textverarbeitungsprogrammen zum Einsatz kommen und fester Bestandteil des Microsoft-Office-Paketes sein [4].

Die deutschen Hochschulen sind dabei, für sich Definitionen im Umgang mit dem Einsatz von KI-Sprachtools im Rahmen der Lehre zu finden [5]. Die Einsatzpalette von Sprach-KI kann dabei von Lehr- und Lernunterstützung über Automatisierung und Entlastung von Routinetätigkeiten bis hin zum Umgang mit Sprach-KI bei der Leistungsbemessung reichen. Lehrende stehen vor der Herausforderung Leistungen der Studierenden künftig auch daraufhin zu prüfen, ob ein Text gänzlich oder mit Unterstützung durch Sprach-KI verfasst worden ist. [5]. Allein das Wissen um die Existenz und Möglichkeiten von ChatGPT lässt jeden Text potentiell verdächtig erscheinen. Auch der Einfluss von Sprach-KI auf wissenschaftliche Publikationen wird in Zukunft wachsen. Hier hat Springer Nature einen ersten Schritt getan, indem sie in ihren Autorenrichtlinien vorgeben, dass Sprach-KI keine Autorenschaft übernehmen kann und die Verwendung im Methodenteil oder in der Danksagung angezeigt werden muss [6].

Bei den administrativen Fragen zu Sprach-KI und ChatGPT sollte aber auch die Frage des Nutzens für ein Fachgebiet im Vordergrund stehen. Ist das Programm in der Lage, auf komplexe Fragestellungen einzugehen und Sachverhalte richtig darzustellen? Die Textgestaltung durch ChatGPT erfolgt auf Grundlage einer statistischen Wahrscheinlichkeit, dass ein verwendeter Begriff auf ein anderes Wort folgt. Durch systematisches Training wird die Sprachfähigkeit trainiert und damit verbessert. Aber, ob die antrainierte Sprachfähigkeit auch einen Wahrheitsgehalt hat, das kann ChatGPT nicht erkennen [1].

Erprobung von ChatGPT für medizinbibliothekarische Fragestellungen (April und Mai 2023)

Wir haben aus unterschiedlichen Themenbereichen des medizinischen Bibliotheks- und Informationswesens einige Fragestellungen in ChatGPT eingegeben und einer Ergebnisprüfung unterzogen. Die angeführten 8 Anwendungsbeispiele zeigen beeindruckende Ergebnisse, aber auch Grenzen von ChatGPT in der Version 3.5 auf.

Beispiel 1: Therapieempfehlung und medizinische Fachliteratur

*Bitte gib eine detaillierte, deutschsprachige Therapieempfehlung für Epistaxis (Nasenbluten) bei Erwachsenen. Gib die Inhalte in Form einer Punktaufzählung an. Zielgruppe sind junge Ärzt*innen, die Unterstützung bei Behandlungsentscheidungen suchen. Gib jeweils auch Verweise auf die Fachliteratur an.*

Abbildung 1 zeigt einen Screenshot der Antwort von ChatGPT.

Beurteilung: Die Antwort von ChatGPT auf diese behandlungsbezogene, medizinische Fragestellung enthält relevante Punkte, bleibt aber trotz gegenteiliger Aufforderung oberflächlich. Auch ohne medizinische Expertise kann die Antwort mit spezialisierten Systemen wie etwa AMBOSS bei weitem nicht mithalten.

Bibliothekarisch spannend sind die Literaturverweise, von denen nur der erste korrekt ist, sowohl im Hinblick darauf, dass der Artikel existiert, als auch auf seine thematische Ausrichtung. Die anderen sind dagegen erfunden bzw. zusammengewürfelt. Das Erfinden von Inhalten ist ein bekannter Kritikpunkt an ChatGPT, der die Anwendung im wissenschaftlichen oder bibliothekarischen Kontext aktuell problematisch macht. Aktuelle Entwicklungen zeigen hier eine Verbesserung und möglicherweise werden fehlerhafte Literaturverweise irgendwann der Vergangenheit angehören – im Vergleich mit wesentlich „weicheren“ Informationen sollte das technisch beherrsch-

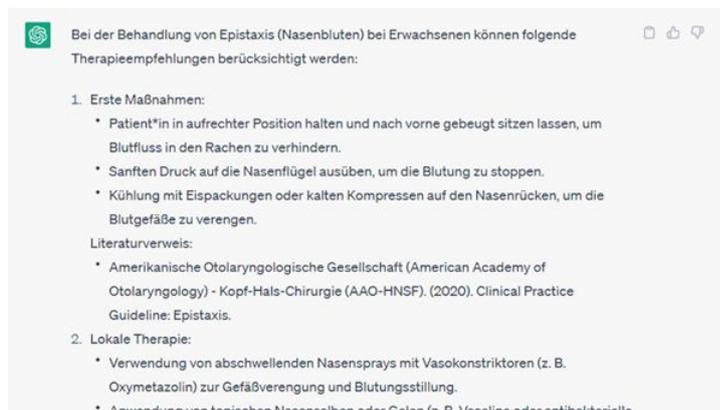


Abbildung 1: Therapieempfehlung und medizinische Fachliteratur

bar sein. Allerdings zeigen diese „Kinderkrankheiten“ doch, wie sehr die ausgegebenen Inhalte aus unklaren Quellen zusammengewürfelt sind und dass die Basis ihrer Glaubwürdigkeit sehr stark die guten sprachlichen Formulierungsfähigkeiten der KI sind. Ob solche Systeme tatsächlich Erkenntnisse aus der Gesamtheit publizierter Fachliteratur zu einem Thema generieren können, könnte weniger ein technisches als ein Problem des Zugangs zu den Fachartikeln sein, die ja weiterhin großteils „closed access“ sind. Naheliegender sind eher Systeme, die Antworten aufgrund einer geschlossenen Textbasis erzeugen, also etwa ein UpToDate-Bot, mit dem Ärzt*innen ggf. auch sprechen könnten.

Beispiel 2: Berufsbild

Medizinische Bibliothekare arbeiten in einem dynamischen beruflichen Umfeld. Bitte stelle die Herausforderungen der nächsten Jahre dar und gib eine Übersicht, wie sich medizinische Bibliothekare hier professionell positionieren und weiterentwickeln können.

Abbildung 2 zeigt einen Screenshot der Antwort von ChatGPT.

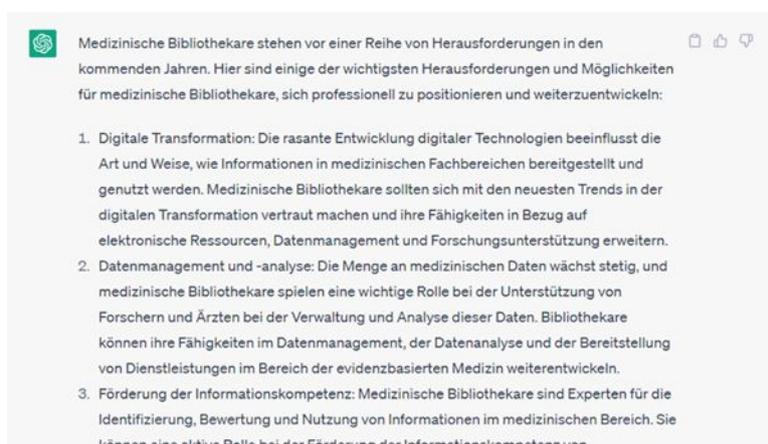


Abbildung 2: Berufsbild

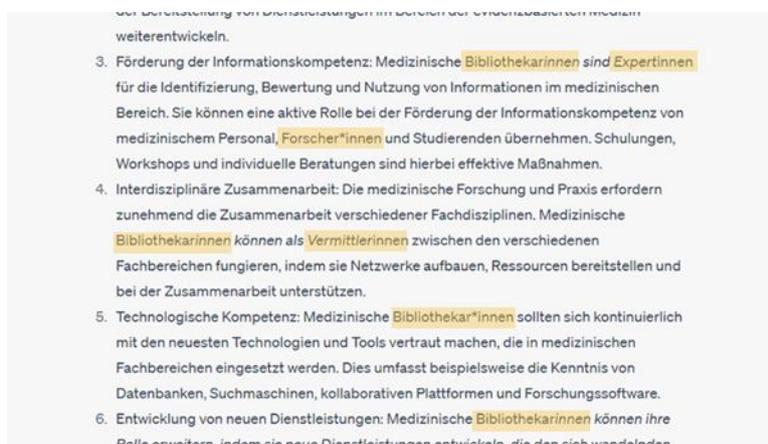


Abbildung 3: Berufsbild gendern

Bitte gendere die Antwort noch mit Sternchen.

Abbildung 3 zeigt einen Screenshot der Antwort von ChatGPT.

Beurteilung: Die Darstellung der professionellen Situation von medizinischen Bibliothekar*innen ist sehr treffend und gut formuliert. Die Ausdrucksweise ist etwas „trocken“ und könnte mehr Details vertragen. Stilistische Verbesserungen sind möglicherweise bereits jetzt mit einer Optimierung der Fragestellung (Prompting) möglich und ganz sicher Gegenstand von Weiterentwicklungen in naher Zukunft.

Die Aufforderung, den Text noch mit “*” zu gendern, hat ChatGPT dann nur halbkorrekt absolviert und nur manche Begriffe mit dem Sternchen versehen – grundsätzlich wurde der Text allerdings gendert.

Beispiel 3: Datenbankrecherche mit PubMed

Schreibe mir einen Searchstring für PubMed, um nach dem Thema Diabetes Mellitus Typ 2 und unblutige Blutzuckermessung zu suchen.

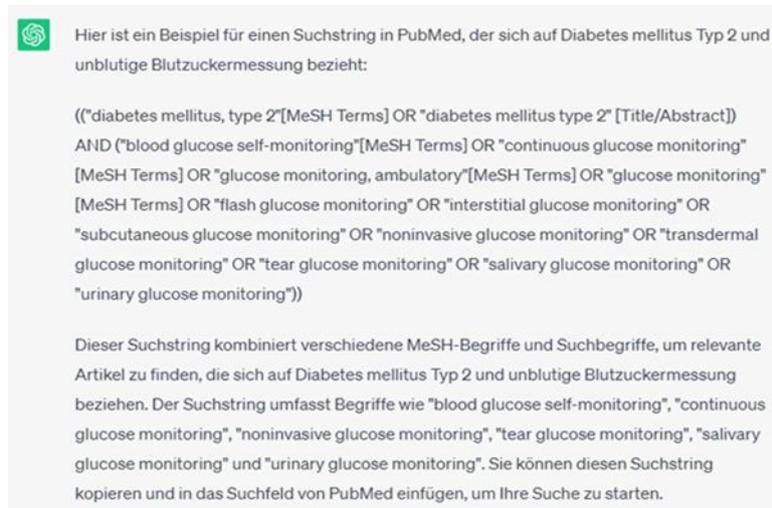


Abbildung 4: Datenbankrecherche mit PubMed

Abbildung 4 zeigt einen Screenshot der Antwort von ChatGPT.

Beurteilung: Die Suche scheint auf den ersten Blick plausibel und ist mit 2 korrekt geklammerten Suchblöcken logisch aufgebaut. Die Bool'schen Operatoren wurden richtig angewandt und der für das direkte Einfügen in PubMed aufbereitete Suchstring funktioniert in PubMed tadellos.

Das Thema Diabetes Mellitus Typ 2 wird korrekt, aber wenig variantenreich abgebildet. Neben dem richtigen MeSH-Begriff wären weitere Stichwort-Varianten – und nicht nur der 1:1 übernommene MeSH-Term – wünschenswert gewesen. Interessant ist, dass die für systematische Suche empfohlene Kombination der Suchfelder MeSH-Term und Title/Abstract verwendet wird.

Ganz anders beim Thema unblutige Blutzuckermessung. Neben dem korrekten MeSH-Begriff "blood glucose self-monitoring"[MeSH Terms] werden einfach weitere MeSH-Begriffe frei erfunden ("continuous glucose monitoring"[MeSH Terms], "glucose monitoring, ambulatory"[MeSH Terms], "glucose monitoring"[MeSH Terms]). Diese könnten als Stichworte durchaus interessante Aspekte zur Suche beitragen. Es folgt eine variantenreiche Liste mit Stichworten als Suchstring, die jetzt aber nicht mehr im Suchfeld „Title/Abstract“ sondern ohne Bezeichnung – also in All Fields – gesucht werden.

Im abschließenden Absatz der ChatGPT-Antwort folgt neben der sinnvollen Erklärung der Methode („Dieser Suchstring kombiniert verschiedene MeSH-Begriffe und Suchbegriffe“) und dem Hinweis auf das Einfügen in PubMed noch eine redundante Aufzählung von einzelnen Stichwörtern ohne Relevanz.

Durch die freie Erfindung von MeSH-Begriffen und unscharfe Verwendung der Suchfelder wird der Suchstring insgesamt unbrauchbar. Nichtsdestotrotz wäre die ChatGPT-Antwort ein guter Ausgangspunkt, um die eigene Suche aufzubauen. Die vielen Varianten der Suchbegriffe können zu einer umfassenderen Suchstrategie anregen.

Beispiel 4: Aufgabe zur Suche in PubMed erstellen

Stelle eine Aufgabe zur Suche mit Filtern in PubMed.

Abbildung 5 zeigt einen Screenshot der Antwort von ChatGPT.

Beurteilung: Die bewusst sehr offen gestellte Frage führt zu einer interessanten Aufgabenstellung in PubMed, wobei ChatGPT sowohl ein interessantes medizinisches Thema als auch die zu filternden Aspekte „systematische Übersichtsarbeiten“ und 5 Jahre Publikationszeitraum benennt. Das dann allerdings noch erwähnte „Peer-Review-Verfahren“ ist redundant.

Die Schritt-für-Schritt Anleitung enthält viele Fehler, den Button „Filter“ gibt es in PubMed ebenso wenig wie den Filter „Custom range“ (der sich natürlich intuitiv erschließen lässt) oder die Journal Category „Peer-reviewed journals“. Auch der in der Aufgabestellung geforderte Filter systematische Übersichtsarbeiten wird zu „Review“ verkürzt.

ChatGPT schlägt vor, nach Literatur der letzten 5 Jahre zu suchen. In diesem Zusammenhang könnte es irritierend wirken, dass in einer Anfrage aus dem Jahr 2023 auf die Jahre 2016-2021 verwiesen wird. Auf Grund der nur bis 2021 reichenden Datenbasis von ChatGPT in der Version 3.5 ist dies jedoch konsequent.

Die Antwort kann allenfalls als Anregung für die eigene Aufgabenstellung gesehen werden, die konkrete Umsetzung in PubMed ist zu einem Großteil frei erfunden.

 Aufgabe: Sie sind auf der Suche nach einer systematischen Übersichtsarbeit zum Thema "Diabetes und kardiovaskuläre Erkrankungen", die in den letzten fünf Jahren veröffentlicht wurde und in einer Fachzeitschrift mit Peer-Review-Verfahren veröffentlicht wurde. Verwenden Sie die Filteroptionen von PubMed, um die Suchergebnisse zu begrenzen und die relevantesten Artikel zu finden.

Schritte:

1. Öffnen Sie PubMed und geben Sie "diabetes cardiovascular disease" in das Suchfeld ein.
2. Klicken Sie auf "Search", um die Suchergebnisse anzuzeigen.
3. Klicken Sie auf "Filter" unter der Suchleiste.
4. Wählen Sie "Custom range" und geben Sie "2016" als Startdatum und "2021" als Enddatum ein, um die Suche auf die letzten fünf Jahre zu beschränken.
5. Wählen Sie unter "Article types" die Option "Review" aus, um nur systematische Übersichtsarbeiten zu erhalten.
6. Unter "Language" wählen Sie die Sprache, die Sie bevorzugen.
7. Unter "Additional filters" wählen Sie "Journal categories" und wählen Sie "Peer-reviewed journals" aus, um sicherzustellen, dass nur Artikel aus Fachzeitschriften mit Peer-Review-Verfahren angezeigt werden.
8. Klicken Sie auf "Show" und überprüfen Sie die Suchergebnisse, um sicherzustellen, dass sie Ihren Kriterien entsprechen.
9. Überprüfen Sie die Titel und Abstracts der Artikel und wählen Sie die relevantesten aus.

Durch die Verwendung der Filteroptionen in PubMed können Sie die Suchergebnisse auf die relevantesten Artikel beschränken und Zeit sparen, indem Sie gezielt nach Artikeln suchen, die Ihren spezifischen Anforderungen entsprechen.

Abbildung 5: Aufgabe zur Suche in PubMed erstellen

 Um eine Webseite im Vancouver-Stil zu zitieren, müssen Sie folgende Elemente angeben:

1. Autorenname (Nachname und Initialen, wenn vorhanden).
2. Titel der Webseite oder des Artikels.
3. Name der Website (falls unterschiedlich vom Titel).
4. URL der Webseite.
5. Datum des Zugriffs auf die Webseite (in der Reihenfolge Jahr-Monat-Tag).

Das Format für das Zitieren einer Webseite im Vancouver-Stil lautet wie folgt:

Autor(en). Titel der Webseite oder des Artikels [Internet]. Name der Website. Datum der Veröffentlichung. Verfügbar unter: URL [Zugriffsdatum].

Beispiel:

Smith J. How to Cite a Website in Vancouver Style [Internet]. Citation Machine. 2021. Verfügbar unter: <https://www.citationmachine.net/vancouver/cite-a-website> [Zugriffsdatum: 12. April 2023].

Bitte beachten Sie, dass das Format je nach den spezifischen Anforderungen Ihrer Institution oder Publikation variieren kann. Es ist immer eine gute Idee, die genauen Anforderungen zu überprüfen, um sicherzustellen, dass Ihre Zitate korrekt formatiert sind.

Abbildung 6: Zitieren nach Vancouver

Beispiel 5: Zitieren nach Vancouver

Wie zitiere ich eine Webseite nach Vancouver?

Abbildung 6 zeigt einen Screenshot der Antwort von ChatGPT.

Beurteilung: Die Antwort wird als gut bewertet, allenfalls sollte noch darauf hingewiesen werden, dass Institutionen

bei Webseiten eine wichtige Rolle spielen und ggf. ebenfalls zitiert werden sollten. Die dann als Beispiel zitierte Website gibt es wirklich und steht auch in inhaltlichem Zusammenhang mit der Frage.

Der Disclaimer am Schluss wird als sehr positiv bewertet. Die KI könnte im Themenfeld Zitierstile bereits zum Generieren von E-Mail-Antworten auf Anfragen verwendet werden – mit nur minimalem Korrekturaufwand.

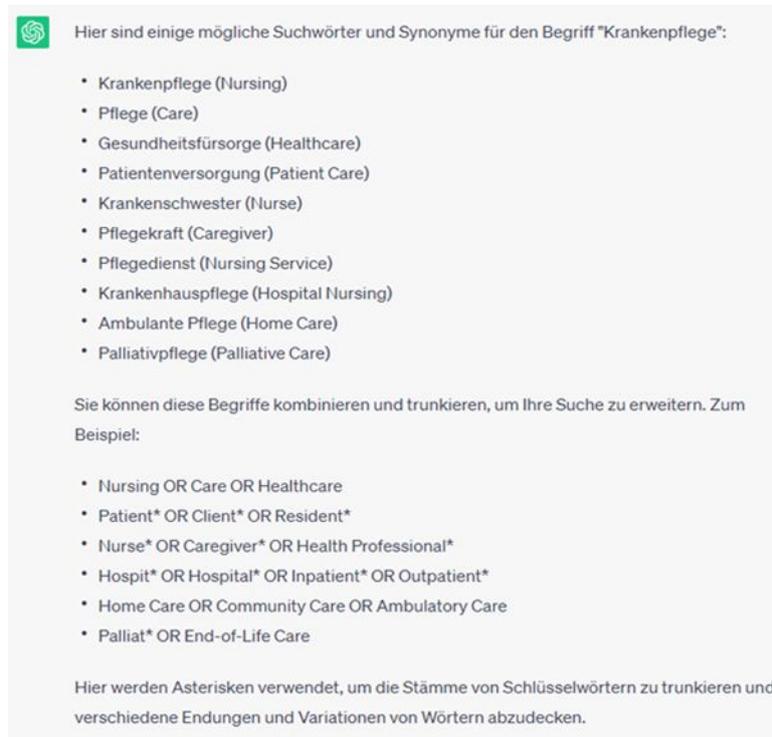


Abbildung 7: Suchwörter und Synonyme

Beispiel 6: Suchwörter und Synonyme

Welche Suchwörter und Synonyme können für den Begriff Krankenpflege verwendet werden. Wie könnte trunkiert werden?

Abbildung 7 zeigt einen Screenshot der Antwort von ChatGPT.

Beurteilung: ChatGPT liefert verschiedene Suchwörter und Synonyme zum Aspekt Krankenpflege auf Deutsch und Englisch. Diese Begriffe können direkt für eine Suchstrategie übernommen und getestet werden.

Bei der Frage nach der Trunkierung fällt auf, dass nur einige Begriffe korrekt aus dem ersten Teil der Antwort übernommen werden. Andere zusammengesetzte Begriffe werden aus dem Zusammenhang gerissen und gehen damit thematisch in eine falsche Richtung (Hospital, Patient). Die Trunkierungen wirken sehr pauschal und sind in der Praxis wenig hilfreich (z. B. Nurse* – besser wäre Nurs* für Nurse, Nursing etc.).

Nicht unerwähnt bleiben darf die von ChatGPT verwendete schöne Bezeichnung „Asterisken“ für das Trunkierungssternchen.

Es zeigt sich erneut, dass die KI zum jetzigen Zeitpunkt für eine Ideensammlung gut geeignet, hingegen für die komplexe Bearbeitung – wie z.B. die Trunkierung – nur sehr eingeschränkt verwertbar ist.

Beispiel 7: Literaturlisten korrigieren

Bitte bringe nachfolgende Literaturliste in Ordnung und formatiere Sie im APA style. Gib wenn möglich Verweise auf die DOI zu den Artikeln an:

1 Link MS, Berkow LC, Kudenchuk PJ, et al. Part 7: adult advanced

cardiovascular life support: 2015 American Heart Association

Guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. *Circulation* 2015; 132 (suppl 2): S444–64.

2 Yannopoulos D, Bartos JA, Aufderheide TP, et al. The evolving role of

the cardiac catheterization laboratory in the management of patients

with out-of-hospital cardiac arrest: a scientific statement from the

American Heart Association. *Circulation* 2019; 139: e530–52.

3 Bartos JA, Grunau B, Carlson C, et al. Improved survival with

extracorporeal cardiopulmonary resuscitation despite progressive

metabolic derangement associated with prolonged resuscitation.

Circulation 2020; 141: 877–86.

4 Grunau B, Reynolds J, Scheuermeyer F, et al. Relationship between

time-to-ROSC and survival in out-of-hospital cardiac arrest ECPR

candidates: when is the best time to consider transport to hospital?

Prehosp Emerg Care 2016; 20: 615–22.

5 Grunau B, Reynolds JC, Scheuermeyer FX, et al. Comparing the

prognosis of those with initial shockable and non-shockable

rhythms with increasing durations of CPR: informing minimum

durations of resuscitation. *Resuscitation 2016; 101: 50–56.*

6 Yannopoulos D, Bartos JA, Raveendran G, et al. Coronary artery

disease in patients with out-of-hospital refractory ventricular

fibrillation cardiac arrest. *J Am Coll Cardiol 2017; 70: 1109–17.*

7 Lamhaut L, Tea V, Raphalen JH, et al. Coronary lesions in refractory

out of hospital cardiac arrest (OHCA) treated by extra corporeal

pulmonary resuscitation (ECPR). *Resuscitation 2017; 126: 154–59.*

Articles

1816 www.thelancet.com Vol 396 December 5, 2020

Abbildung 8 zeigt einen Screenshot der Antwort von ChatGPT.

Bitte formatiere die Liste nun im Bibtext-Format zum Import in ein Literaturverwaltungsprogramm.

Abbildung 9 zeigt einen Screenshot der Antwort von ChatGPT.

Beurteilung: Eingegeben wurde eine Literaturliste, deren Formatierung beim Kopieren aus einem PDF-Dokument beschädigt wurde, zum Beispiel wurden fehlerhafte Zeilenumbrüche eingefügt. ChatGPT korrigiert und transformiert die Literaturliste weitgehend korrekt in einen anderen Zitationsstil. Bei den Seitenzahlen ergänzt ChatGPT sogar die abgekürzte Schreibweise, z.B. “e530-52” wird zu “e530-e532”. Die zusätzlich angeforderten DOIs (Digital Object Identifier) sind allerdings nicht korrekt, entweder sind sie ungültig oder sie verweisen auf einen anderen Text. Auf „Textreste” aus der Fußzeile, die aus dem PDF mitkopiert wurden, wird hingegen korrekt hingewiesen. Diese Leistung ist ebenso beeindruckend wie hilfreich, insbesondere, wenn man davon ausgeht, dass die „Erfindungsfreudigkeit“ des Systems in zukünftigen Versionen sicherlich noch eingeschränkt wird und damit die Vertrauenswürdigkeit der Antworten steigen wird.



Abbildung 8: Korrigierte Literaturliste

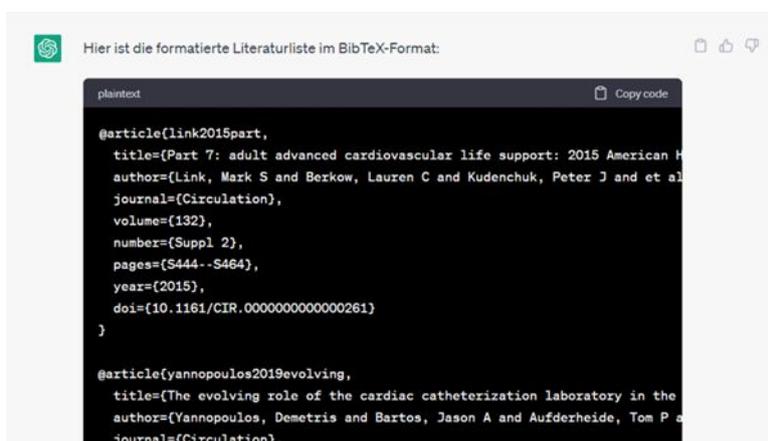


Abbildung 9: Literaturliste als Bibtex

In einer weiteren Fragestellung wurde ChatGPT gebeten, die Literaturliste in das BibTex-Format zum Import in Literaturverwaltungsprogramme zu formatieren, das macht das System ziemlich perfekt.

Beispiel 8: Dublettenprüfung in EndNote

Schreibe eine Anleitung für die Durchführung einer Endnote Dublettenprüfung.

Abbildung 10 zeigt einen Screenshot der Antwort von ChatGPT.

Beurteilung: ChatGPT hat anscheinend nicht erkannt, dass sich die Anfrage auf das Literaturverwaltungsprogramm EndNote und dessen Dublettenprüfung bezogen hat. Aus diesem Grund geht die Antwort deutlich am Thema vorbei und enthält inhaltlich nicht schlüssige, bestenfalls allgemein gültige Überlegungen zur Dublettenkontrolle.

Eine Konkretisierung der Fragestellung mit Benennung der Arbeitsschritte im EndNote-Menü ergibt ein wesentlich besseres Ergebnis:

Schreibe eine Anleitung für eine Endnote Dublettenprüfung mit den Arbeitsschritten Edit, Preferences, duplicates, library, find duplicates.

Abbildung 11 zeigt einen Screenshot der Antwort von ChatGPT.

Beurteilung: Mit der konkretisierten Fragestellung hat ChatGPT eine Anleitung mit „fast“ exakter Benennung der nötigen Arbeitsschritte verfasst. Die Abfolge der Arbeitsschritte sind in richtiger Reihenfolge mit korrekter Bezeichnung der Menüpunkte erfolgt. Einige Sätze in der Anleitung sind jedoch in ihrer Aussage nicht korrekt, so z.B. die Aussage „ob Sie Dubletten in der gleichen Bibliothek oder auch in anderen Bibliotheken finden möchten. Passen Sie die Einstellungen nach Ihren Bedürfnissen an“. Eine zeitgleiche Dublettenprüfung in mehreren parallel geöffneten EndNote-Dateien ist nicht möglich. Die Dublettenprüfung muss in jeder Datei einzeln vorgenommen werden.

In den Schritten 5 und 6 wird die Durchführung der Dublettenprüfung beschrieben, ohne jedoch konkret auf die Abläufe einzugehen. Es fehlt der Hinweis, dass sich ein neuer Tab öffnet, in dem die identifizierten Dubletten zum Vergleich gegenübergestellt werden. Auch auf die individuelle Auswahlmöglichkeit (keep this Record) in dieser Gegenüberstellung wird nicht eingegangen. Die im weiteren empfohlene Zusammenführung von Dubletten gibt es nur im Literaturverwaltungsprogramm Zotero. EndNote kennt diese Funktion nicht.

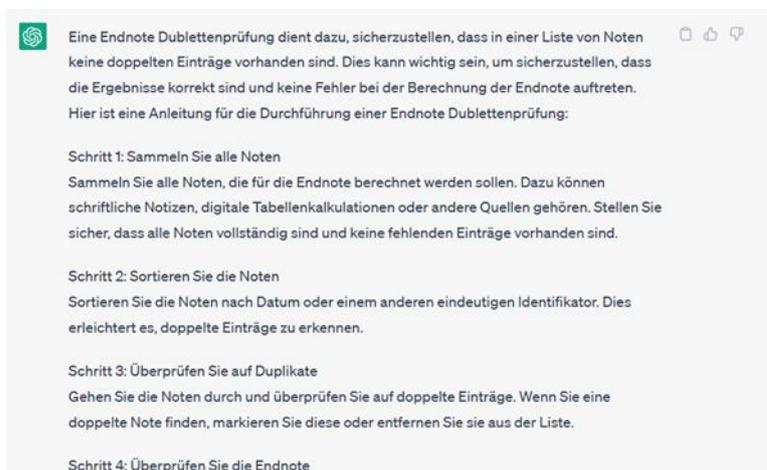


Abbildung 10: EndNote Dublettenprüfung

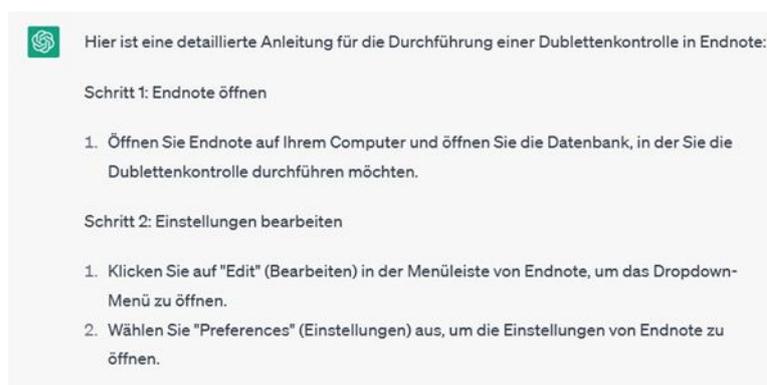


Abbildung 11: Dublettenprüfung in EndNote – Detailangaben

Anleitungen zum Handling von Softwareprogrammen müssen detaillierte Abläufe darstellen und sollten Arbeitsschritte nicht auslassen oder überspringen. Ansonsten droht die Gefahr zu verwirren. Hier zeigt ChatGPT Schwächen. Gänzlich falsch ist die Antwort von ChatGPT nicht, jedoch lässt sie vermuten, dass ChatGPT bei der Antwort auch auf Texte anderer Literaturverwaltungsprogramme zugegriffen hat. Zwar ist die Reihenfolge in der Menüführung korrekt beschrieben, aber auf die Beschreibung der Durchführung zur Dublettenprüfung verzichtet ChatGPT gänzlich. Es mag daran liegen, dass es bei Literaturverwaltungsprogrammen wie Citavi oder Zotero keinen nachvollziehbaren Arbeitsprozess zur Dublettenprüfung gibt. ChatGPT war anscheinend nicht in der Lage, die gesammelten Texte der unterschiedlichen Literaturverwaltungsprogramme voneinander zu trennen.

Fazit und Ausblick

Auch wenn die behandelten Fragen nur eine winzige Momentaufnahme aus einem sich rasch wandelnden Bereich geben, sind die Antworten von ChatGPT doch gleichzeitig beeindruckend wie auch irritierend. Sie zeigen zum einen das enorme Potenzial auf, das derartige Systeme haben, und es kann ohne Übertreibung angenommen werden, dass hier in kurzer Zeit noch erhebliche Weiterentwicklungen zu erwarten sind.

Zum anderen sind die Antworten aber auch immer wieder fehlerhaft und oft auf unerwartete Weise, die sich zwar technisch mit der rein textbasierten, assoziativ-statistischen Arbeitsweise von ChatGPT erklären lässt, aber doch auch auf unsere Erwartungen an ein vorgeblich „intelligentes“ System verweist. Die Fehler sind nicht in der von uns erwarteten Weise „dumm“, sondern vielfach unverständlich und wirken geradezu provokant.

ChatGPT schreibt beeindruckende Texte in dem Sinne, dass tatsächlich neue und grammatikalisch einwandfreie Sprache produziert wird, sowohl auf Englisch als auch auf Deutsch und vermutlich auch in weiteren Sprachen. Bringt man das System durch entsprechende „Prompts“ auf die richtige Spur, sind die Ergebnisse häufig korrekte Sachverhaltsdarstellungen, Berichte, Zusammenfassungen etc. Korrektheit ist für ChatGPT aber kein Kriterium an sich – falsche Inhalte und kreative Erfindungen werden weithin kritisiert – vielmehr wird ein oberflächlich glaubwürdiger Text auf Basis riesiger vorher analysierter Textmengen generiert, manchmal eben ohne und manchmal mit Fehlern, wie sich auch in den von uns gestellten Fragen erwiesen hat. Allgemeine Fragen zum Berufsbild medizinischer Bibliothekar*innen werden tadellos im Stil einer „Sonntagsrede“ beantwortet, auch medizinische Behandlungsfragen sind meist im Wesentlichen korrekt, solange die Antwort an der Oberfläche bleibt, ebenso Antworten zur Literaturrecherche, Bedienung von Softwareprodukten etc. Wenn es konkret und detailliert wird, kommt es aber vermehrt zu Fehlern, etwa in der genauen Vorgangsweise bei der Benutzung eines (bestimmten!) Literaturverwaltungsprogramms oder wenn es um Litera-

turverweise geht. Bei letzterem ist ChatGPT besonders erfindungsfreudig, generiert beliebige DOIs oder würfelt Metadaten durcheinander – was aber letztendlich nur unseren Erwartungen an „Intelligenz“ widerspricht, vermutlich nicht aber den Vorstellungen der KI von einem gültigen Antworttext.

An diesen Problemen wird sich noch einiges ändern, vieles wird besser werden, vieles ist schon verfügbar und wurde aus pragmatischen Gründen in diesem Artikel ausgeklammert, um angesichts der rasanten Entwicklung überhaupt eine solche „Momentaufnahme“ zu ermöglichen. Der Abgleich der Fragen mit aktuellen Ergebnissen einer Internetsuche ist zum Beispiel bereits möglich, ebenso entstehen gerade vielerlei KI-Angebote für spezielle Zwecke, z.B. die Literaturrecherche, oder es werden unterschiedliche KI-Systeme kombiniert, die nicht – wie ChatGPT nur auf Sprachanalysen beruhen.

Auch die Anwendung der KI auf eine definierte Wissensbasis birgt große Chancen, z.B. um Fragen über eine umfassende Website zu beantworten oder um schnell Informationen aus einem klinischen Informationssystem wie UpToDate zu generieren.

Insofern werden viele oberflächliche Fehlleistungen wohl bald der Vergangenheit angehören. Was wohl bleiben wird, ist die Frage nach der Glaubwürdigkeit der Ergebnisse, die sogar umso drängender wird, wenn die Ergebnisse nicht mehr „mit bloßem Auge“ als richtig oder falsch zu erkennen sind. Noch spannender wird diese Entwicklung, wenn man bedenkt, dass diese Systeme Antworten zukünftig in großer Menge produzieren werden, so dass „händisch“ produzierte Inhalte zurückgedrängt werden bzw. nur als neue „Inspiration“ für die Verarbeitung durch KI-Systeme dienen werden, die dann wieder computer-generierte Varianten davon erzeugt. Es wäre wünschenswert, dass ChatGPT und andere KI ihre Texte direkt mit einer Quellenangabe versehen, die auch verdeutlicht, dass der Text nicht reproduziert werden kann, Beispiel: „Dieser Text wurde am 19.05.2023 von ChatGPT generiert und ist nicht reproduzierbar. URL.“. Ansonsten ist die Quellenangabe selbstverständlich von den Personen anzugeben, die diese Texte verwenden. Die Ansprache von Chat-Bots wie ChatGPT scheint auf die Ergebnisse zwar keinen wesentlichen Einfluss zu haben, könnte aber durchaus große Auswirkungen auf die menschliche Bewertung der Antwort haben. Ob die KI menschlich angesprochen wird, ob vertraut mit Du oder distanziert mit Sie, ob unpersönlich sachlich oder in einem Befehlstone setzt die Menschen in eine Beziehung mit der KI. Deshalb sollte dieser Faktor sorgfältig abgewogen werden. Um sich die aktuellen Unschärfen bewusst zu machen, könnte die KI mit einer Formulierung wie „Mache Vorschläge zu ...“ angesprochen werden. So ist klar, dass es sich nicht um einen belastbaren Text, sondern wirklich „nur“ um eine brillant formulierte – aber ungeprüfte – Antwort handelt.

Viele und vor allem arbeitsintensive, repetitive Arbeiten mit Daten aller Art werden in Bälde von KI-Assistenzsystemen übernommen werden, zum Teil von spezialisierten Tools, zum Teil wohl auch von allgemeinen Assisten-

ten, die dann vielleicht (endlich!) leisten, was die Apologeten der Digitalisierung schon lange versprochen haben. Die Auswirkungen auf die beruflichen Aufgaben auch im Bereich des Medizinischen Bibliothekswesens sind dabei unabsehbar, bereits in der Vergangenheit kam es durch technische Veränderungen ja auch immer wieder zum Entstehen neuer Aufgabenbereiche und Tätigkeiten. Als Experten im Informationsbereich können sich wissenschaftliche Bibliothekar*innen aber durchaus an die Spitze einer Bewegung setzen, die Veränderungen im Bereich der Wissensproduktion und -verbreitung durch KI-Systeme begleitet und hoffentlich auch gestaltet. Sehr offen und fragil sind aber die Fragen, die über akademische Abläufe und Traditionen hinausweisen und wo KI-Systeme gesellschaftliche Prozesse beeinflussen werden, wie etwa digitale Überwachung, Fake-News, Fragen echter und künstlicher Identitäten etc. Das einzige, das an dieser Stelle als sicher erachtet werden kann, ist, dass eine Entwicklung mit weitreichenden Folgen gerade erst begonnen hat. Hoffen wir, dass wir die positiven Chancen auch genießen können!

Anmerkung

Interessenkonflikte

Die Autoren erklären, dass sie keine Interessenkonflikte in Zusammenhang mit diesem Artikel haben.

Literatur

- Spitzer M. ChatGPT. *Nervenheilkunde*. 2023;42(04):192-9. DOI: 10.1055/a-1948-8785
- Schmohl T, Watanabe A, Schelling K, Hrsg. Künstliche Intelligenz in der Hochschulbildung: Chancen und Grenzen des KI-gestützten Lernens und Lehrens. Bielefeld: transcript; 2023. (Hochschulbildung: Lehre und Forschung; Band 4). DOI: 10.25656/01:26427
- Schick U. Was ist künstliche Intelligenz? 20.03.2018 [zuletzt aufgerufen: 03.08.2023]. Verfügbar unter: <https://news.sap.com/germany/2018/03/was-ist-kuenstliche-intelligenz/>
- Lindner R. ChatGPT: Microsoft bringt Open AI ins Büro. *Frankfurter Allgemeine Zeitung*. 16.03.2023. Verfügbar unter: <https://www.faz.net/aktuell/wirtschaft/chatgpt-microsoft-bringt-open-ai-in-word-outlook-und-powerpoint-18754133.html>
- Salden P, Leschke J, Hrsg. Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung. Zentrum für Wissenschaftsdidaktik der Ruhr-Universität Bochum; 2023. DOI: 10.13154/294-9734
- Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. *Nature*. 2023;613(7945):612. DOI: 10.1038/d41586-023-00191-1

Korrespondenzadresse:

Dr. Stefanus Schweizer
 Universitätsbibliothek Mainz, Bereichsbibliothek
 Medizin, Langenbeckstraße 1, Gebäude 304, 55131
 Mainz, Deutschland
 s.schweizer@ub.uni-mainz.de

Bitte zitieren als

Schweizer S, Vogel C, Weiner O. Können wir künstliche Intelligenzen wie ChatGPT produktiv für die medizinbibliothekarische Arbeit nutzen? *GMS Med Bibl Inf*. 2023;23(1):Doc04. DOI: 10.3205/mbi000554, URN: urn:nbn:de:0183-mbi0005546

Artikel online frei zugänglich unter

<https://doi.org/10.3205/mbi000554>

Veröffentlicht: 13.09.2023

Copyright

©2023 Schweizer et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.