

Einsatz von Large Language Models in der Suchtberatung: Inhaltsanalyse der Antwortqualität von ChatGPT

Marieke Theil^{1*}

¹ Delphi Gesellschaft für Forschung, Beratung und Projektentwicklung mbH, Berlin, Deutschland

* Korrespondenz, E-Mail: marieke.theil@delphi.de

© 2025 Marieke Theil; Lizenznehmer Infinite Science Publishing

Dies ist ein Open-Access-Abstract, das unter den Bedingungen der Creative Commons Attribution License veröffentlicht wird, welche uneingeschränkte Nutzung, Verbreitung und Reproduktion in jedem Medium erlaubt, sofern das Originalwerk ordnungsgemäß zitiert wird. (<http://creativecommons.org/licenses/by/4.0>)

Hintergrund und Fragestellung

Der potenzielle Einsatz von Large Language Models (LLMs) im Suchtberatungskontext gewinnt zunehmend an Relevanz. Vor diesem Hintergrund ist es essenziell, den Inhalt und die Qualität LLM-generierter Antworten zu bewerten sowie mögliche damit einhergehender Grenzen und Risiken kritisch zu reflektieren. Die Untersuchung analysiert, inwieweit ChatGPT in seinen Antworten zentrale Wirkfaktoren der Online-Beratung umsetzt, wie die Antwortqualität zu bewerten ist und welche Unterschiede sich hinsichtlich dessen zwischen Antworten mit und ohne vorgeschalteten Prompt mit Instruktionen zum Antwortverhalten zeigen.

Methoden

Zur Untersuchung wurden 15 prototypische Beratungsanfragen formuliert, die jeweils einmal ohne und einmal mit einem spezifischen Prompt von ChatGPT beantwortet wurden. Die generierten Antworten wurden mithilfe der inhaltlich-strukturierenden und evaluativen qualitativen Inhaltsanalyse nach Kuckartz & Rädiker (2022) untersucht, um die Anwendung und Umsetzung relevanter Wirkfaktoren sowie bestehende Limitationen und Risiken systematisch zu erfassen.

Ergebnisse

Die Ergebnisse zeigen, dass ChatGPT primär Wirkfaktoren der Bewältigungsperspektive integriert, insbesondere durch Handlungsempfehlungen und die Weiterleitung an professionelle Stellen. Auch einfache Reflexionen wurden häufig identifiziert, während gezielte Nachfragen nur selten vorkamen. Die Mehrheit der Antworten wurde als fallbezogen bewertet. Ein vorgeschalteter Prompt ging mit einer verbesserten Bewertung einzelner Wirkfaktoren einher, etwa der Analyse von Beratungszielen, der Verwendung permissiver Sprache, der Problembewältigung, der Triage, der Anteilnahme und der Toleranz. Etwa die Hälfte der Antworten wies sachliche Ungenauigkeiten auf, während potenziell risikobehaftete Aussagen nur vereinzelt auftraten.

Diskussion und Schlussfolgerung

Die Analyse zeigt, dass die generierten Antworten Parallelen zu professionellen Beratungspraktiken aufweisen. Dennoch weisen die Ergebnisse darauf hin, dass sich das Modell als unzureichend für komplexe, interaktive Beratungsprozesse erweist. Die Ergebnisse unterstreichen die Notwendigkeit weiterführender Forschung, insbesondere dem Einfluss der Optimierung von LLMs durch spezifische Trainingsstrategien und adaptive Prompt-Designs.

Offenlegung von Interessenskonflikten sowie Förderungen

Während der letzten 3 Jahre bestanden folgende wirtschaftlichen Vorteile oder persönliche Verbindungen, die die Arbeit zum eingereichten Abstract beeinflusst haben könnten: Die Autorin ist beruflich bei der Delphi Gesellschaft für Forschung, Beratung und Projektentwicklung mbH tätig, die verantwortlich ist für das vom Bundesministerium für Gesundheit (BMG) geförderte Projekt SuchtGPT (Gestaltung, Programmierung und Testung eines KI-basierten Chatbots für Suchtfragen). Die hier vorgestellten Ergebnisse der Masterarbeit entstanden jedoch unabhängig davon und basieren

ausschließlich auf einer eigenständig durchgeführten Untersuchung im Rahmen des Masterstudiengangs M.Sc. Prävention und Gesundheitspsychologie an der SRH Fernhochschule.

Erklärung zur Finanzierung: Der Beitrag basiert auf einer Masterarbeit, die im Rahmen eines eigenfinanzierten Studiums an der SRH Fernhochschule unter der Betreuung von Prof. Dr. Daniele Ackermann-Piek entstand.

