

## Medizinische Angaben gegenüber KI sind oft lückenhaft

**Eine neue Studie zeigt: Schildern Menschen ihre Krankheitssymptome einer KI statt ärztlichem Fachpersonal, sinkt die Qualität ihrer Angaben. Dies gefährdet die Genauigkeit digitaler Diagnosen und die Patientensicherheit.**

Gut möglich, dass in naher Zukunft Menschen einer KI ihre Symptome schildern müssen, bevor sie einen Arzttermin erhalten. Die KI entscheidet dann, ob es sich um einen Notfall handelt oder ob die Behandlung noch warten kann, und vergibt dementsprechend Termine.

So weit ist es glücklicherweise noch nicht, doch auch im Gesundheitswesen schreitet die Digitalisierung rasant voran. KI-Chatbots und digitale Symptom-Checker übernehmen eine immer wichtigere Rolle und dienen zunehmend als erste Anlaufstelle für die sogenannte „Self-Triage“ – also die erste Einschätzung der Dringlichkeit einer Behandlung durch die Betroffenen selbst.

Doch während die technische Leistungsfähigkeit dieser Systeme stetig wächst, rückt ein anderer Faktor in den Fokus der Forschung: das menschliche Kommunikationsverhalten im Dialog mit der Maschine. Denn selbst die beste Technik, insbesondere in der medizinischen Diagnostik, ist auf präzise Informationen angewiesen, die Nutzende jedoch nicht immer im vollen Umfang bereitstellen.

### **Menschliche Zurückhaltung bremst das Potenzial der KI aus**

Das ist das zentrale Ergebnis einer jetzt in der Fachzeitschrift Nature Health veröffentlichten Studie. Verantwortlich dafür sind Professor Wilfried Kunde, Inhaber des Lehrstuhls für Psychologie III der Julius-Maximilians-Universität Würzburg (JMU), und Moritz Reis, Wissenschaftlicher Mitarbeiter an diesem Lehrstuhl. Daran beteiligt waren Wissenschaftlerinnen und Wissenschaftler der Charité – Universitätsmedizin Berlin, der University of Cambridge sowie des Helios Klinikums Emil von Behring und vom Vivantes Klinikum Neukölln in Berlin.

„Die 500 Studienteilnehmerinnen und -teilnehmer hatten die Aufgabe, simulierte Symptombereiche für zwei weit verbreitete Krankheitsbilder – ungewöhnliche Kopfschmerzen und Grippe – zu verfassen“, beschreibt Erstautor Moritz Reis das Studien-Design. Dabei wurden sie in dem Glauben gelassen, dass ihre Berichte entweder von einem KI-Chatbot oder einem menschlichen Arzt gelesen würden. Ziel war es, die Qualität dieser Berichte im Hinblick auf ihre Eignung für eine medizinische Dringlichkeitseinschätzung zu untersuchen.

### **Qualitätsverlust zeigt sich in geringerer Detailtiefe**

Die zentrale Erkenntnis: Glaubten die Teilnehmenden, mit einer Künstlichen Intelligenz zu kommunizieren, verschlechterte sich die Eignung ihrer Schilderungen für eine medizinische Ersteinschätzung messbar im Vergleich zur Interaktion mit vermeintlichem ärztlichem Fachpersonal. Dieser Effekt war sogar bei jenen Teilnehmenden zu sehen, die zum Zeitpunkt der Befragung tatsächlich unter den entsprechenden Symptomen litten.

Dieser Qualitätsverlust zeigt sich direkt in der Detailtiefe der Berichte. Während die Beschreibungen gegenüber Fachkräften im Durchschnitt 255,6 Zeichen umfassten, waren es bei Chatbots lediglich 228,7 Zeichen.

Auch wenn ein Unterschied von 28 Zeichen nach wenig klingt, ist er nach Aussagen des Forschungsteams von Bedeutung und kann zur Folge haben, dass selbst hochleistungsfähige KI-Modelle am Ende falsche medizinische Ratschläge geben. Schließlich scheitern auch diese an der medizinischen Einordnung, wenn Patientinnen und Patienten nicht alle wesentlichen Informationen mitteilen. Der Erfolg digitaler Erstberatung hängt nämlich weniger von der Rechenleistung als von der menschlichen Bereitschaft zur detaillierten Schilderung.

### **Psychologische Barrieren: Die Sorge vor der „Einheitsdiagnose“**

Doch warum verhalten sich Menschen gegenüber Maschinen so zurückhaltend? Ein wesentlicher Grund ist vermutlich die sogenannte „Vernachlässigung der Einzigartigkeit“ (Uniqueness Neglect). „Viele Menschen nehmen an, dass eine KI die individuellen Besonderheiten ihrer persönlichen Situation nicht erfassen kann und stattdessen lediglich standardisierte Muster abgleicht“, erklärt Wilfried Kunde.

Zusätzlich führen Skepsis gegenüber der Diagnosefähigkeit von Algorithmen sowie Datenschutzbedenken dazu, dass Angaben oft abgekürzt oder vage gehalten werden. Moritz Reis bringt die menschliche Komponente so auf den Punkt: „Wenn wir einer Maschine nicht zutrauen, unsere Einzigartigkeit zu verstehen, enthalten wir ihr unbewusst die Informationen vor, die sie für eine präzise Hilfe bräuchte.“ Dieser psychologische Filter sorgt dafür, dass medizinisch relevante Details gar nicht erst im System ankommen und die Diagnosequalität sinkt.

### **Den Dialog mit der Maschine verbessern**

Die Forschungsergebnisse zeigen nach Ansicht des Forschungsteams deutlich, dass die technische Weiterentwicklung der KI allein nicht ausreicht. Eine mögliche Lösung sehen sie deshalb in einer klugen Gestaltung der Benutzeroberflächen.

Um die Qualität der Symptomberichte zu steigern, sollten Entwicklerinnen und Entwickler konkrete Beispiele für hochwertige Beschreibungen bereitstellen und die KI so programmieren, dass sie fehlende Details aktiv und gezielt nachfragt. Nur wenn Nutzende dazu ermutigt werden, reichhaltige Informationen zu liefern, können Fehlbehandlungen vermieden und das Gesundheitssystem effektiv entlastet werden.

### **Originalpublikation:**

Reduced Symptom Reporting Quality During Human-Chatbot Versus Human-Physician Interactions, Moritz Reis, Florian Reis, Yeun Joon Kim, Aylin Demir, Jess Lim, Matthias I. Gröschel, Sebastian D. Boie, Wilfried Kunde. Nature Health, DOI: 10.1038/s44360-026-00116-y