

Was die neue WhatsApp-KI kann und welche Gefahren sie birgt

Vertrauliche Chats, kreative Posts, smarte Antworten: Der neue Helfer in den Meta-Apps bringt Komfort und wirft neue ethische Fragen auf. Drei Experten der Universität Luxemburg erklären

Von Sebastian Weisbrodt

Ein neuer Ring sorgt derzeit für Gesprächsstoff in sozialen Netzwerken: Ein blau-roter Kreis, der plötzlich in Apps wie WhatsApp, Facebook oder Instagram auftaucht. Dahinter verbirgt sich MetaAI, der neue KI-Assistent des gleichnamigen US-amerikanischen Konzerns, der nun auch in Europa ausgerollt wurde. Doch was kann die Technologie wirklich, wo liegen die Chancen, und welche Risiken sind nicht zu unterschätzen?

Laut Meta markiert die neue Funktion einen Meilenstein in der Nutzung künstlicher Intelligenz im Alltag. Noch nie war ein Sprachmodell so tief in die sozialen Netzwerke integriert – und noch nie war eines für Nutzerinnen und Nutzer so leicht zugänglich. Der Zuwachs an digitaler Funktionalität geht allerdings mit einer Herausforderung in Sachen Medienkompetenz, Datenschutz und digitaler Ethik einher.

„MetaAI kann man als einen alltagstauglichen, kreativen Begleiter beschreiben. Die KI ist nicht nur ein reiner Lieferant von Information“, sagt Professor Christoph Schommer, Experte für Künstliche Intelligenz an der Universität Luxemburg. Für weniger technikaffine Nutzer bietet sich damit ein einfacher Einstieg in die Welt der KI – ohne technologische Hürden.

Ein Assistent in der Hosentasche

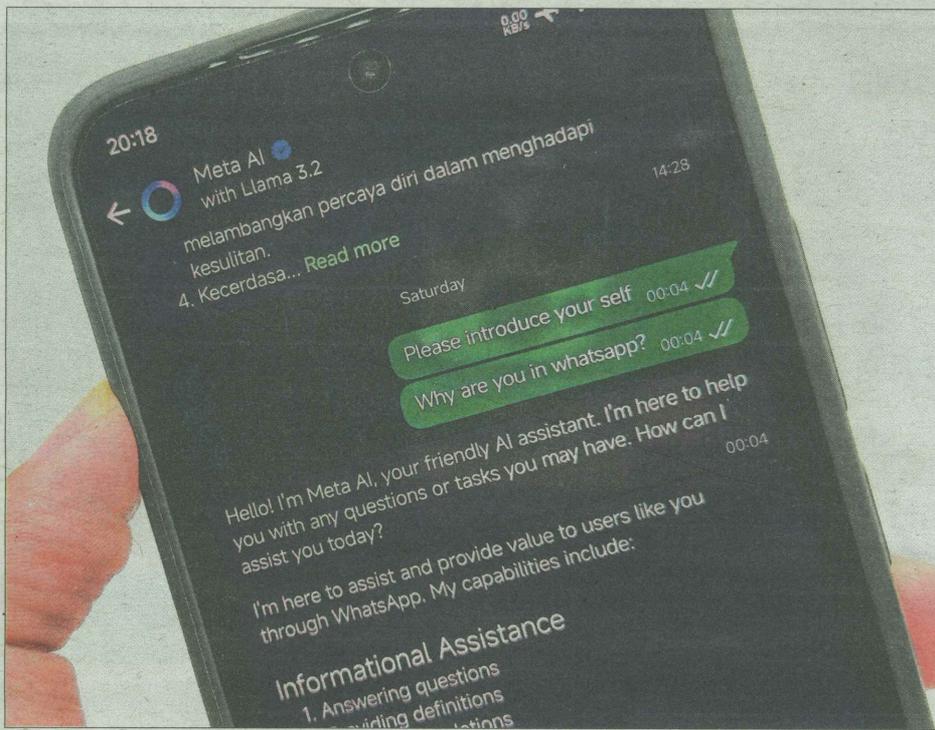
MetaAI basiert auf dem hauseigenen Sprachmodell LLaMA 3 (Large Language Model Meta AI 3), einem Open-Source-Modell, das auf Milliarden von Parametern trainiert wurde. Anders als klassische Chatbots wie ChatGPT oder Google Gemini ist MetaAI nicht als eigenständige App, sondern direkt in die Plattformen des Meta-Konzerns eingebettet. Ein Fingertipp auf den neuen farbigen Ring genügt, und schon öffnet sich ein Chatfenster, in dem Nutzer Fragen stellen, Texte formulieren oder kreative Inhalte generieren lassen können.

Der Vorteil: Die Anwendung ist nahtlos, niedrigschwellig und durchgehend verfügbar – ein digitaler Begleiter, der ständig mit-schwingt, ohne zusätzliche Programme oder Apps zu benötigen. Genau darin sehen viele Fachleute jedoch auch eine Gefahr: Die Nähe zur privaten Kommunikation verwischt die Grenzen zwischen Mensch und Maschine.

Alltagshelfer mit Potenzial

Meta verspricht eine breite Funktionspalette: von schnellen Alltagsantworten über Übersetzungen bis hin zu Rezepttipps oder Hilfe beim Verfassen von Social-Media-Posts. In Instagram können – perspektivisch – sogar Bilder aus Textanweisungen generiert werden. In der Europäischen Union steht aufgrund der strengen Datenschutzregeln bislang allerdings nur die Textfunktion zur Verfügung.

In WhatsApp agiert MetaAI wie ein persönlicher Assistent: Er hilft beim Formulieren von Nachrichten, bei der Übersetzung in Fremdsprachen oder beantwortet spontane Fragen wie „Was ist ein gutes Urlaubsziel für eine 80-Jährige?“ oder „Was ist ein guter Countrysong?“ Auf Facebook unterstützt die KI beim Schreiben von Posts, liefert Ideen für Kommentare oder sogar ganze Gruppenbeschreibungen. Nutzer erkennen dabei stets, dass sie mit einer



Meta hat seine KI direkt in Facebook, Instagram und Whatsapp eingebunden. Foto: Shutterstock

KI kommunizieren – doch die Übergänge wirken oft fließend.

Ein weiteres Beispiel: Kleine Unternehmen oder Selbstständige können MetaAI nutzen, um Beiträge für ihre Firmenprofile zu formulieren, Produkte zu bewerben oder Kundenanfragen schneller zu beantworten. Eltern wiederum erhalten Unterstützung beim Planen von Geburtstagsfeiern, beim Erklären von Hausaufgaben oder beim Finden altersgerechter Inhalte.

Wie funktioniert das eigentlich?

Technisch gesehen basiert MetaAI auf der sogenannten Transformer-Architektur. Die

WhatsApp: Kann man Meta AI löschen oder deaktivieren?

Derzeit bietet WhatsApp keine offizielle Möglichkeit, Meta AI vollständig zu löschen oder zu deaktivieren. Die KI wird jedoch nur aktiv, wenn sie gezielt aufgerufen wird. Wer die Funktion nicht nutzen möchte, kann sie einfach ignorieren. Es gibt aber einige Tricks, um den Assistenten zumindest aus dem Blickfeld zu entfernen.

Wenn man Meta AI nicht verwenden möchte, bieten sich folgende Wege an:

- Nicht aktivieren: Meta AI reagiert ausschließlich, wenn sie über den blauen Button oder durch eine direkte Ansprache mit „@Meta AI“ in Gruppenchats aktiviert wird. Solange man die KI nicht anspricht, bleibt sie passiv.
- Chat archivieren oder löschen: Falls der Meta-AI-Chat in der Übersicht stört, kann man ihn archivieren, indem man den Chat nach links wischt und die Option „Archivieren“ auswählt. Alternativ lässt sich der Chat auch löschen: Dazu wischt man ebenfalls nach links, tippt auf „Mehr“ (bei Android) beziehungsweise direkt auf „Löschen“ (bei iOS), und bestätigt den Vorgang.

Auch wenn Meta AI dadurch nicht vollständig aus WhatsApp entfernt wird, verschwindet sie zumindest optisch aus der Chatübersicht – und bleibt so aus dem alltäglichen Blickfeld.

dritte Generation des hauseigenen LLaMA-Modells verarbeitet riesige Mengen an Text aus öffentlich verfügbaren Quellen wie Wikipedia, Online-Foren oder Nachrichtenportalen. Das Modell erkennt Muster, lernt Sprachzusammenhänge und kann so passende Antworten erzeugen – in Sekundenschnelle. In der Praxis bedeutet das: Gibt man MetaAI eine Frage oder eine Aufgabe, berechnet es aus Milliarden Trainingsbeispielen die wahrscheinlichste Antwort.

Meta betont, dass dabei keine privaten Nachrichten ausgelesen würden – zumindest nicht in WhatsApp. Anders sieht es bei öffentlichen Facebook-Posts aus, die in anonymisierter Form durchaus zum Training herangezogen werden können. Die Inhalte, die Nutzer im Chat mit MetaAI eingeben, werden genutzt, um den Gesprächsverlauf zu verstehen, aber nicht dauerhaft gespeichert – es sei denn, die Nutzer stimmen dem explizit zu.

Zwischen Effizienz und Eingriff in die Privatsphäre

Doch so bequem der neue KI-Begleiter auch ist – die Einführung erfolgt nicht ohne Kritik. Datenschutz- und Verbraucherschützer mahnen zur Vorsicht. Zwar betont Meta, dass private Chats in WhatsApp nicht ausgelesen würden, öffentliche Inhalte auf Facebook oder Instagram jedoch sehr wohl zur Verbesserung der KI beitragen könnten. Auch die Prompts – also das, was Nutzer der KI schreiben – werden analysiert, um bessere Antworten zu geben.

Für die Informatikerin und Sprachwissenschaftlerin Dr. Salima Lamsiyah, ebenfalls von der Universität Luxemburg, ist genau hier ein kritischer Punkt erreicht: „Viele Nutzer wissen gar nicht, welche Daten sie der KI übermitteln – und was daraus abgeleitet werden kann. Transparenz ist dringend nötig, nicht nur technische Raffinesse.“ Sie fordert verpflichtende digitale Aufklärung – insbesondere für jüngere Nutzer, die KI-Anwendungen als selbstverständlich erleben, aber deren Mechanismen kaum verstehen.

Ähnliche Töne schlägt Professor Schom-

mer, assoziiertes Mitglied am Luxembourg Institute of Digital Ethics, an: „Die Integration von MetaAI in Kommunikationsapps bedeutet, dass zwischenmenschliche Gespräche nun potenziell mitgehört, analysiert und sogar beeinflusst werden können – auch wenn das nur indirekt geschieht. Der Schutz der Privatsphäre wird dadurch fragiler.“ Besonders problematisch sei laut Schommer die „erzwungene KI“ – Nutzer können die Funktion nicht vollständig deaktivieren, sondern lediglich ignorieren. „Das widerspricht dem Prinzip der informierten, selbstbestimmten Nutzung“, so Schommer weiter.

Soziale und sprachliche Aspekte

Meta betont, es gebe umfangreiche Sicherheitsmaßnahmen: Inhaltsfilter, Feedbacksysteme, Quellprüfung. Nicht-überprüfbare, problematische oder gefährliche Inhalte würden automatisch gefiltert. Doch das Phänomen der sogenannten „Halluzinationen“ – also falsche oder frei erfundene Antworten – ist auch bei MetaAI nicht verschwunden. „Das System überprüft sich selbst – aber eben nur teilweise. So kann es zu Fehlinformationen aufgrund einer unzureichenden oder fehlerhaften Datenlage sowie nicht identifizierter Fake News kommen“, sagt Aria Nourbakhsh, Experte für Chatbot-Technologie an der Universität Luxemburg. Die Folgen seien weitreichend, da daraus Diskriminierungen, Vertrauensverluste, Manipulation oder Cybermobbing entstehen könnten. „Die Gefahr ist, dass Menschen den Aussagen der KI mehr Glauben schenken, als sie sollten. Gerade weil sie so nahtlos in bekannte Apps eingebettet ist“, so Nourbakhsh weiter.

Für ihn ist MetaAI ein Beispiel für ein größeres Dilemma: „Wir haben es mit Systemen zu tun, die Menschen in Echtzeit begleiten, aber dabei nicht transparent zeigen, woher Informationen stammen oder welche Interessen dahinter stehen.“ Es brauche klare ethische Leitlinien.

Auch die gesellschaftlichen Folgen bleiben nicht aus. Schommer spricht von einer „Verarmung der natürlichen Sprache“: „Je mehr wir mit KI-Systemen kommunizieren, die effizient, formelhaft und kontextlos antworten und je mehr wir uns auf diesen digitalen Diener verlassen, desto stärker verlieren wir auch ein Stück Identität. Und eine sprachliche Vielfalt, Ironie oder kulturelle Nuancen könnten so auf der Strecke bleiben.“

Kommunikation auf dem Prüfstand

Die breite Einführung von MetaAI wirft Fragen nach der Zukunft der Kommunikation auf. Prof. Schommer sieht eine gesellschaftliche Pflicht zur Reflexion: „Wir müssen nicht nur fragen, was die Technik kann, sondern auch, was wir wollen. Wollen wir wirklich, dass Künstliche Intelligenz unser Sprachverhalten mitprägt?“

Die drei Forscher meinen, dass man sich im Leben etwas bewahren muss: „Dazu gehört die Privatsphäre genauso wie der Datenschutz und eben eine intellektuelle Selbstfürsorge. Wir lehnen den Einbau von KI als festen Bestandteil einer Kommunikations-App prinzipiell ab, auch wenn diese hilft und eine Funktionalität durch Ein- und Ausschalten geregelt werden kann. Stattdessen schlagen wir diese als Option vor, die nur bei explizitem Ausdruck eines Willens eingeschaltet werden kann.“