

Traumatische Klickarbeit: Die Menschen hinter ChatGPT

Von Anne-Careen Stoltze (BFH Wirtschaft) | 0 Kommentare



Die Fähigkeiten neuer generativer KI-Modelle oder darauf basierende Produkte wie ChatGPT sind faszinierend. Neben den Bedenken, was dies mit unserer Gesellschaft macht, kommt mehr und mehr Kritik auf an den Arbeitsbedingungen der beteiligten Menschen in solchen Projekten. Wir haben darüber mit unseren Expertinnen gesprochen.

SocietyByte: Was ist dran an der Kritik an der Aufgabe der Arbeiter*innen? Was mussten sie machen?

Mascha Kurpicz-Briki: Zur Lancierung von ChatGPT wurde Anfang 2023 eine Recherche vom TIME Magazine veröffentlicht [1]. Diese zeigte, dass das US-amerikanische Unternehmen OpenAI, dem Chat GPT gehört, eine Firma in Kenia beauftragte, um toxische Inhalte wie körperliche und sexuelle Gewalt, Suizide und Tierquälerei, aus den Antworten von zu filtern. Dabei mussten die angestellten Arbeitskräfte für weniger als 2 Dollar die Stunde teils schockierende Inhalte lesen. Gemäss den Recherchen lasen die Arbeiter*innen in einer neunstündigen Schicht bis zu 250 Textpassagen von einer Länge bis zu 1.000 Wörtern. Es wurde auch kritisiert, dass die betroffenen Mitarbeitenden zu wenig Betreuung im Umgang damit erhielten, sie beklagten psychische Probleme aufgrund der belastenden Texte.

Wofür wurde ihre Vorarbeit genutzt bzw. wofür diente sie?

MKB: Auf Grund der grossen Menge von Trainingsdaten, die für solche Sprachmodelle benötigt werden, ist eine Qualitätskontrolle schwierig. Der Inhalt der Trainingsdaten kann daher dazu führen, dass diskriminierende oder beleidigende Aussagen von einem Chatbot generiert werden. Um das zu verhindern, können solche Antworten beispielsweise markiert werden als unerwünscht. Wenn man dies bei sehr vielen Antworten macht, so kann das System davon lernen. Um es markieren zu können, muss man jedoch all diese unerwünschten Texte lesen, und dies kann unter anderem detaillierte Beschreibungen von Missbrauch, Folter oder Mord enthalten.

Wie kommt es, dass so eine Filterung überhaupt notwendig ist?

MKB: Das Grundproblem liegt darin, dass auf Grund der grossen Mengen an Trainingsdaten, die für solche Modelle erforderlich sind, Daten aus dem Internet verwendet werden. Die Menge der Daten ist so gross, dass eine manuelle Kontrolle der Daten gar nicht mehr möglich ist, jedoch enthalten diese teilweise sehr problematische Daten aus den ganz dunklen Ecken des Internets. Um diese automatisch zu erkennen, und zu verhindern, dass dies in der Antwort vom Chatbot verwendet wird, braucht es wiederum eine grosse Menge an Beispielen von solchen schlimmen Texten, von denen das System lernen kann, was unerwünscht ist. Diese Texte müssen durch Menschen bereitgestellt und sortiert werden.

Warum wurde nicht bereits in der Trainingsphase inklusiver, d.h. reflektierter in Bezug auf Rassismus und Bias, gearbeitet?

MKB: Die Daten, welche für das Training verwendet werden, wurden durch Menschen generiert. Sie enthalten daher auch die Stereotypen unserer Gesellschaft. In der Trainingsphase werden die Daten automatisch verarbeitet, und wir reden von einer Grössenordnung von Hunderten von Milliarden von Wörtern. Die Auswahl an hier verfügbaren Trainingssätzen ist daher beschränkt. Die Qualität der Trainingsdaten ist sehr wichtig, und es wird immer mehr gefordert, dass dies ausführlicher und transparenter dokumentiert wird [2].

Ist das ein spezifisches Problem von OpenAI oder ChatGPT?

MKB: Es gibt diese Arten von Tätigkeit im Rahmen von KI und der digitalen Transformation auch an anderer Stelle. Auch bei der Filterung von Inhalten auf den sozialen Medien, oder in Online-Foren braucht es oft Menschen. In einigen Fällen, um direkt eine Inhaltsprüfung durchzuführen, oder aber um KI-Systeme zu trainieren, die dies dann machen sollen. Wie genau es dort jeweils um die Arbeitsbedingungen steht, insbesondere wenn ein Outsourcing in den globalen Süden erfolgt, ist natürlich schwierig pauschal abzuschätzen.

Neben dieser Problematik gibt es gerade im Bereich Sprachmodelle noch weitere Herausforderungen. Viele Fortschritte werden primär für die englische Sprache, oder für einige wenige privilegierte Sprachen, erzielt. Ein Grossteil der Sprachen, die weltweit gesprochen werden, kann nicht profitieren davon, da sich die Forschung und Entwicklung weniger damit auseinandersetzt.

Warum wird diese Arbeit ausgelagert und so wenig gewertschätzt?

Caroline Straub: Niemand in der Schweiz könnte sein Leben mit dieser Arbeit finanzieren. Es handelt sich um microwork. Das sind einfache, repetitive Aufgaben, die schnell online erledigt werden können und keine besonderen Fähigkeiten erfordern (z.B. Datenreinigung, Coding, Klassifizierung von Kontent für die KI). Die Bezahlung für Mikrojobs ist in der Regel sehr niedrig (ca. 5 Rappen pro Click). Für viele Menschen ohne Bildung in Entwicklungsländern (Globaler Süden) sind microjobs eine Möglichkeit, Geld zu verdienen. Microwork wird auch als Ghostwork (Geisterarbeit) bezeichnet. Das ist Arbeit, die von einem Menschen ausgeführt wird, von der ein Kunde jedoch glaubt, dass sie von einem automatisierten Prozess ausgeführt wird.

Was sind die Schwierigkeiten bezüglich den Anstellungsbedingungen, mit Ausweitung auf Ghostwork?

CS: Künstliche Intelligenz ist auf menschliche Arbeit angewiesen, um Aufgaben wie Datenbereinigung, Codierung und Klassifizierung von Inhalten auszuführen. Diese On-Demand-Arbeit wird online auf Plattformen wie Amazon Mechanical Turk angeboten und ausgeführt, wobei die Bezahlung je nach Aufgabe erfolgt. Diese als „Geisterarbeit“ konzipierte, schnell wachsende, plattformbasierte Arbeit ist weitgehend unsichtbar: Arbeitnehmer*innen können nicht mit Manager*innen sprechen, erhalten kein Feedback und es mangelt an Arbeitsschutz.

Referenzen

[1] <https://time.com/6247678/openai-chatgpt-kenya-workers/>

[<https://time.com/6247678/openai-chatgpt-kenya-workers/>]

[2] [https://cacm.acm.org/magazines/2021/12/256932-datasheets-for-](https://cacm.acm.org/magazines/2021/12/256932-datasheets-for-datasets/abstract)

[datasheets/abstract](https://cacm.acm.org/magazines/2021/12/256932-datasheets-for-datasets/abstract) [[https://cacm.acm.org/magazines/2021/12/256932-](https://cacm.acm.org/magazines/2021/12/256932-datasheets-for-datasets/abstract)

[datasheets-for-datasets/abstract](https://cacm.acm.org/magazines/2021/12/256932-datasheets-for-datasets/abstract)]

Über die Expertinnen



Mascha Kurpicz-Briki ist Professorin für Data Engineering an der Berner Fachhochschule und stv. Leiterin der Forschungsgruppe Applied Machine Intelligence. Sie erforscht, wie KI verantwortungsvoll genutzt werden kann.



Prof. Dr. Caroline Straub ist Professorin am Institut New Work der BFH Wirtschaft. Sie forscht zu Plattformbasierter Arbeit, Digital HRM und Diversity & Inclusion.



AUTHOR: ANNE-CAREEN STOLTZE



Anne-Caren Stoltze ist Redaktorin des Wissenschaftsmagazins SocietyByte und Host des Podcasts "Let's Talk Business". Sie arbeitet in der Kommunikation der BFH Wirtschaft, sie ist Journalistin und Geologin.

Posts from Anne-Caren Stoltze | Website

Create PDF

Ähnliche Beiträge

Auch Algorithmen diskriminieren - wie es ihnen ihre Programmierer vorgeben

ChatGPT: Orakel, Doktor oder Joker? - Ethische Grenzen der Nutzung verschiedener Tools

Hi ChatGPT, hast du Vorurteile?

Wenn ChatGPT bei der Prüfung hilft - eine Podcast-Episode über KI im Studium

Wenn Mehmet und Peter nicht gleich sind - Vorurteile auf Grund der Namensherkunft in

Wortvektoren

Wie Bias in KI-Systemen erkannt und verringert werden können

«Wir müssen Bias aus Sprachmodellen entfernen» - eine Podcastfolge über KI in Verwaltung und Justiz

0

COMMENTS