

W05_Ethik_Benachteiligung und blinde Flecken

Benachteiligung und blinde Flecken

Erarbeitet von

Marie von Lobenstein M.A.

Lernziele	1
Inhalt	1
Einstieg	1
Was ist Klickarbeit?.....	2
Beispiele für Klickarbeit	2
Ethische Implikationen von Klickarbeit.....	3
Take-Home Message	4
Quellen	4
Disclaimer	5

Lernziele

- Du weißt was mit „blinden Flecken“ in Bezug auf KI-Systeme gemeint ist
- Du kannst erklären, was Klickarbeit ist
- Du kannst ein Beispiel für Klickarbeit nennen
- Du kannst ethische Implikationen von Klickarbeit nennen

Inhalt

Einstieg

KI-Ethik konzentriert sich häufig auf "Bedenken" verschiedener Art, was eine typische Reaktion auf neue Technologien ist. Der Buchdruck hat die Menschen in Aufruhr gebracht. Die Erfindung des Telefons beispielsweise ließ viele den Untergang der Gesellschaft fürchten, weil sie Angst hatten, dass Menschen nicht mehr persönlich miteinander reden.

Diese Ängste mögen heute für manche albern sein, aber zugrunde liegt ihnen etwas, was wir alle kennen; die Angst vor Neuem. So ist es auch mit KI-Technologien. Diese Ängste können dann zu einem Problem werden, sobald sie uns blind für reale Risiken werden lassen und uns davon ablenken eine informierte Einschätzung über die Potentiale und Risiken von KI-Systemen zu treffen. Die unreflektierte Angst vor den zukünftigen Auswirkungen von künstlicher Intelligenz, macht uns blind für Risiken im hier und jetzt. Kate Crawford und Ryan Calo haben schon 2016 auf diese „blinden Flecken“ aufmerksam gemacht. Ich möchte heute mit euch einen aktuellen blinde Flecken, also Risiken, die bislang noch wenig diskutiert werden, besprechen: Die sogenannte Klickarbeit.

Quelle [1][2][3]

Was ist Klickarbeit?

KI-Algorithmen werden nicht, wie vielfach angenommen oder behauptet, nur durch maschinelle Prozesse trainiert, sondern häufig manuell durch Datenbeschriftung, die von Menschen an virtuellen Fließbändern durchgeführt wird.

Quelle [4]

Clickwork, auch bekannt als "Human-in-the-Loop" oder "Crowdsourced"-Datenannotation, bezieht sich auf die Praxis des Einsatzes menschlicher Arbeitskräfte zur Kennzeichnung und Kategorisierung von Daten für die Verwendung beim Training von KI-Modellen.

Beispiele für Klickarbeit

Ein Beispiel für Clickwork wäre der Einsatz menschlicher Arbeitskräfte zur Kennzeichnung und Kategorisierung von Bildern, um ein KI-Modell zur Identifizierung von Objekten in Bildern zu trainieren.

Bei diesem Verfahren würde man dem menschlichen Arbeiter eine große Anzahl von Bildern zeigen und ihn bitten, bestimmte Objekte auf jedem Bild zu identifizieren und zu beschriften, z. B. Autos, Gebäude oder Fußgänger. Anhand der beschrifteten Daten wird dann ein maschinelles Lernmodell trainiert, um diese Objekte in neuen Bildern zu erkennen und zu klassifizieren.

Du hast um übrigen wahrscheinlich auch schonmal eine KI-Trainiert, unbezahlt wohlgemerkt. Wenn du bei Google mal beweisen solltest, dass du kein Roboter bist, indem du Ampeln in einer Sammlung von Fotos identifizieren solltest, dann hast du eine KI trainiert! Diese Tests werden CAPTCHA genannt, ein Akronym für Completely Automated Public Turing test to tell Computers and Humans Apart.

Ein weiteres Beispiel für Clickwork ist die Kategorisierung von schädlichen Inhalten Social-Media-Plattformen.

Bei diesem Verfahren wird menschlichen Mitarbeiter*innen eine große Anzahl von Bildern, Videos, Beiträgen und Kommentaren mit der Bitte gezeigt, diejenigen zu identifizieren und zu markieren, die Hassreden, Gewaltdarstellungen oder andere schädliche Inhalte enthalten. Die markierten Inhalte werden dann von Moderator*innen überprüft, um zu entscheiden, ob sie von der Plattform entfernt werden sollten. Diese Art von Klickarbeit ist notwendig, um die Sicherheit und das Wohlergehen der Nutzer*innen auf Social-Media-Plattformen zu gewährleisten und dafür zu sorgen, dass auf

Nachrichten-Websites zuverlässige Informationen verfügbar sind. Es muss aber sichergestellt werden, dass die Mitarbeiter*innen, die diese Aufgaben ausführen, ordnungsgemäß geschult sind und dass ihnen klare Leitlinien und Kriterien an die Hand gegeben werden. Außerdem sollte sichergestellt werden, dass sie nicht über einen längeren Zeitraum hinweg schädlichen Inhalten ausgesetzt sind, da dies ihr geistiges und emotionales Wohlbefinden beeinträchtigen könnte.

Ethische Implikationen von Klickarbeit

Damit wir das sicherstellen können, müssen wir aber über einige ethische Implikationen nachdenken:

Faire Entlohnung: Clickworker*innen werden für ihre Arbeit meist sehr schlecht bezahlt, was zu Ausbeutung und finanzieller Not führen kann. Was besonders Bauchschmerzen bereitet, wenn man bedenkt, dass die Clickwork-Industrie das unsichtbare Rückgrat vieler KI-Technologien ist. Der Markt für Datenetikettierungslösungen von Drittanbietern war 2019 1,7 Milliarden Dollar wert und wird bis 2024 voraussichtlich 4,1 Milliarden Dollar erreichen.

Quelle [5][6][7]

Mangelnde Arbeitsplatzsicherheit: Clickworker*innen haben oft nicht das gleiche Maß an Arbeitsplatzsicherheit wie traditionelle Arbeitnehmer*innen, was zu wirtschaftlicher Instabilität und Unsicherheit führen kann. Es gibt sogar Fälle, in denen Gefängnisinsass*innen diese Clickarbeit ausgeführt haben, ohne die richtige Schulung oder das nötige Training, um mit den Inhalten umgehen zu können.

Quelle [8][9]

Bedenken hinsichtlich des Datenschutzes: Clickworker*innen müssen mit sensiblen Informationen, wie z. B. persönlichen Daten, umgehen, ohne dass diese Arbeiter*innen angemessen geschützt oder geschult sind.

Voreingenommenheit: Clickworker*innen können unbeabsichtigt Voreingenommenheit, also Biases, in die Daten einbringen, die sie beschriften, was zu verzerrten KI-Modellen führen kann.

Verantwortung: Da KI-Modelle, die auf der Grundlage von Crowdsourcing-Daten trainiert wurden, bei der Entscheidungsfindung zum Einsatz kommen, muss sichergestellt werden, dass der Prozess der Datenkommentierung transparent und nachvollziehbar ist und wer für die Fehler und Verzerrungen in den Daten verantwortlich ist.

Take-Home Message

Viele machen sich bereits Gedanken darüber, was beispielsweise mit Arbeiter*innen der Logistikindustrie oder mit Taxifahrer*innen passiert, wenn mehr und mehr autonome Fahrzeuge diese Arbeit übernehmen können. Es werden aber bereits jetzt Arbeiter*innen ausgebeutet in der KI-Industrie, die eben noch zu wenig berücksichtigt werden. Ängste der Zukunft dürfen uns also nicht blind machen für Risiken im Jetzt. Es muss sichergestellt werden, dass die Klickarbeit auf ethische und verantwortungsvolle Weise durchgeführt wird, z. B. durch eine faire Entlohnung und Arbeitsplatzsicherheit für die Arbeitnehmer*innen, den Schutz der Privatsphäre der Arbeitnehmer*innen und die Gewährleistung, dass die Daten auf unvoreingenommene Weise gekennzeichnet werden. Auf der anderen Seite kann die Möglichkeit, solche Klickarbeit als Einzelaufträge zu übernehmen und unabhängig vom Ort zu bearbeiten, für manche Personen überhaupt erst die Möglichkeit eröffnen, einen Zusatzverdienst zu erwirtschaften. Klickarbeit wie sie heute bereits praktiziert wird, berührt also sowohl Prinzipien der fairen Arbeit, des Datenschutzes sowie der korrekten Modellbildung und bieten Anlass über entsprechende Regulierungen nachzudenken.

Quellen

- Quelle [1]** Müller, V. (2020). *Ethics of Artificial Intelligence and Robotics* (E. N. Zalta, Ed.; Stanford Encyclopedia of Philosophy(Summer Edition)). <https://plato.stanford.edu/archives/sum2021/entries/ethics-ai/>
- Quelle [2]** Crawford, K., & Calo, R. (2016). There is a blind spot in AI research. *Nature*, 538(7625), 311–313. <https://doi.org/10.1038/538311a>
- Quelle [3]** Hagendorff, T. (2022). Blind spots in AI ethics. *AI and Ethics*, 2(4), 851–867. <https://doi.org/10.1007/s43681-021-00122-8>
- Quelle [4]** Irani, L. (2015). The cultural work of microwork. *New Media & Society*, 17(5), 720–739. <https://doi.org/10.1177/1461444813511926>
- Quelle [5]** Irani, L. (2019). Justice for Data Janitors. In S. Marcus & C. Zaloom (Eds.), *Think in Public* (pp. 23–40). Columbia University Press. <https://doi.org/10.7312/marc19008-003>
- Quelle [6][8]** Casilli, A. A. (2017). Global digital culture| Digital labor studies go global: Toward a digital decolonial turn. *International Journal of Communication*, 11, 21.
- Quelle [7]** Cognilytica, Data Preparation & Labeling for AI 2020, 2020, pp. 1–37. <https://www.cognilytica.com/download/data-preparation-labeling-for-ai-2020-cgr-dlp20/>

Quelle [9] Hao, K. (2019). *An AI startup has found a new source of cheap labor for training algorithms: prisoners*. MIT Technology Review.
<https://www.technologyreview.com/2019/03/29/136262/an-ai-startup-has-found-a-new-source-of-cheap-labor-for-training-algorithms/>

Disclaimer

Transkript zu dem Video „W05_Ethik_Benachteiligung und blinde Flecken“, Marie von Lobenstein.

Dieses Transkript wurde im Rahmen des Projekts ai4all des Heine Center for Artificial Intelligence and Data Science (HeiCAD) an der Heinrich-Heine-Universität Düsseldorf unter der Creative Commons Lizenz [CC-BY](https://creativecommons.org/licenses/by/4.0/) 4.0 veröffentlicht. Ausgenommen von der Lizenz sind die verwendeten Logos, alle in den Quellen ausgewiesenen Fremdmaterialien sowie alle als Quellen gekennzeichneten Elemente.