

Ludwig-Maximilians-Universität München

Fakultät für Kulturwissenschaften

Institut für Ethnologie

Koloniale Kontinuitäten in der KI-Entwicklung

Die Rolle von Online-Plattformen bei der Reproduktion
globaler Ungleichheiten

SoSe24

Aktuelle Fragestellungen

Unter der Leitung von Prof. Dr. Martin Sökefeld

geschrieben von

Felix Keilhack

12077260

Eschenriederstrasse 29, 81249 München

E-Mail: f.keilhack@campus.lmu.de, tel.: +4915223955580

Inhaltsverzeichnis

Einleitung	3
KI-Trainingsdaten von Clickworkern	3
Online-Plattformen	3
Arbeitsaufträge der Clickworker.....	5
Arbeitsverhältnisse der Clickworker	6
Profil der Clickworker.....	7
KI-Entwicklung entlang der globalen Asymmetrie	8
Sprache als Mittel zur Erschließung von Arbeitsmärkten.....	8
Vertikale & Horizontale Arbeitsteilung.....	10
Zynische Ironie der Produktion von Trainingsdaten.....	11
Schluss.....	12
Literatur	13

Koloniale Kontinuitäten in der KI-Entwicklung

Die Rolle von Online-Plattformen bei der Reproduktion globaler Ungleichheiten

Einleitung

Hornborg argumentiert, dass technologischer Fortschritt, der es ermöglicht, Zeit und Raum zu sparen, stets auf Kosten von Arbeitszeit und natürlichem Raum an anderen Orten geht (Hornborg 2014, 122 ff.). Im Kontext der Entwicklung Künstlicher Intelligenzen (KI) war es mir aufgrund begrenzter Datenlage nicht möglich, das Verhältnis zwischen der für die Produktion von Trainingsdaten aufgewendeten Arbeitszeit und der damit erzielten Zeiteinsparung zu ermitteln. Dennoch bietet dieser Gedanke einen wertvollen Ansatzpunkt für meine Arbeit. Er motiviert eine tiefere Untersuchung der ethischen Implikationen hinter den deontologischen Prinzipien und richtet den Fokus auf die Ursprünge der Intelligenz in KI-Systemen.

Aufgrund der mangelnden Transparenz der sozial-strukturellen Arbeitsverhältnisse von Clickworkern untersucht die vorliegende Hausarbeit im Rahmen des Themas der KI-Trainingsdaten, inwieweit die vorwiegend westliche Entwicklung von Künstlicher Intelligenz (KI) auf der Grundlage der (englischen) Sprache von der Wohlstandsasymmetrie zwischen den Ländern des globalen Nordens und Südens profitiert. Die Arbeit beleuchtet dabei deskriptiv verschiedene Missstände im Akquisitionsprozess von Trainingsdaten; das Ergebnis der Arbeit soll zur Reflexion über ethische Fragestellungen entlang der Lieferkette anregen. Insbesondere werden Probleme im KI-Trainingsprozess auf die individuellen Arbeitsverhältnisse der Clickworker, das internationale Machtgefälle und die globale Wohlstandsverteilung zurückgeführt.

KI-Trainingsdaten von Clickworkern

Online-Plattformen

Im Hinblick auf den Durst von KI-Systemen nach Trainingsdaten etablierte sich in den letzten Jahren ein Geschäftsmodell um digitale Plattformen. Dabei unterscheidet man zwischen (a) Online oder Sharing Plattformen, auf denen Gig-Workers ihre Dienste im Rahmen einer Sharing Economy anbieten, wie z.B. Mitfahr- und Lieferdienste, (b) ein Upwork-Plattformen, auf denen remote Freelancer sogenannte Makro-Tasks übernehmen, z.B. Dienstleistungen wie Webdesign oder Programmieren, und (c) Online-Plattformen, die viele sogenannte Mikro-Tasks – kurze, repetitive,

unterteilte Arbeitspakete, wie z.B. das Identifizieren sowie Labeln von Bildern – an eine Masse an selbstständigen Individuen vermittelt (Tubaro und Casilli 2020, 2; Rani und Furrer 2021, 214). Zwar werden die Arbeiten bei (a), (b), (c) online angenommen, doch bleibt die Dienstleistung von (a) an geographische Standorte gebunden (Graham, Hjorth, und Lehdonvirta 2017; Kretschmer 2024), während (b) & (c) lokal unabhängig sind. Der Unterschied zwischen (b) und (c) liegt darin, dass (b) oft einen (langfristigen) Arbeitsvertrag und meist komplexe sowie zeitintensive Dienstleistungen anbieten, während (c) oft kurzfristige Auftragnehmer:innen-Verträge sind (Rani und Furrer 2021, 214; Tubaro und Casilli 2020, 2). Ideen der Skalierung, Flexibilität und Kosteneffektivität für (c) liegen darin, für den Workload eine breite „Armee“ an sogenannten Clickworkern kurzfristig für ein paar Stunden, statt wenige, festangestellte Arbeiter:innen langfristig zu beschäftigen (Tubaro, Le Ludec, und Casilli 2020, 67–68; Rani und Furrer 2021, 213). Diese Hausarbeit beschäftigt sich mit dem Trend (c) Online-Plattformen von westlichen Unternehmen aus den USA und Europa für die Vermittlung von billigen Mikro-Tasks insbesondere an eine breite Masse an gering-qualifizierte Arbeiter:innen in Ländern des globalen Südens outzusourcen.

Online-Plattformen, wie z.B. Microworkers, CrowdFlower, Clickworker, Appen, Amazon Mechanical Turk, ClixSense und viele mehr, vermitteln mit Standort in Ländern des globalen Nordens eine meist internationale Crowd-Größe von jeweils ungefähr 1 bis 10 Millionen Clickworkern (vgl. Tubaro, Le Ludec, und Casilli 2020; Hornuf und Vrankar 2022; Javaid 2024; Grohmann und Fernandes Araujo 2021) an Technologiekonzerne sowie Industrieunternehmen wie Google, Microsoft, Boeing, Ford, Adobe, um ihre KI-Anwendungen trainieren zu lassen (Kretschmer 2024). Die Plattformen wie Remotask, Sama und ScaleAI haben ihren Firmensitz zwar in San Francisco doch operieren in Uganda, Kenia, Indien, Nepal und auf den Philippinen (Taylor 2023; Tan und Cabato 2023; Kretschmer 2024; Perrigo 2023). Nach Angaben der Weltbank macht Plattformarbeit zwischen 4,4-12,5% der globalen Arbeitskraft aus (Miceli u. a. 2024). Diese Clickworker-Industrie als „unsichtbares Rückgrat vieler KI-Technologien“ bildet einen stetig wachsenden Arbeitsmarkt mit einem Wert von 1,7 Milliarden USD in 2019, mit einer Prognose von 4,1 Milliarden USD in 2024 und 17,1 Milliarden USD in 2030 (Hagendorff 2022, 858; Grand View Research 2023).

Obwohl Konzerne aus dem Silicon Valley ein gemeinnütziges Image sowie die Mentalität pflegen, dass der richtige Einsatz von Technologie die Welt verbessert, würden sie bewusst auf billige Arbeitskräfte in Ländern des globalen Südens setzen, mit dem Argument die Arbeitslosigkeit sowie Armut zu bekämpfen (Kretschmer 2024). Auch die Online Plattform Sama, die mehrere Aufträge von OpenAI in Höhe von 200.000 bis knapp 800.000 USD für KI-Trainingsdaten erhielt, portraitiert sich als ethisch korrekt, weil sie nach ihren Angaben mehr als 50.000 Personen aus der Armut verholpen hätte, mit dem Bezahlen durchschnittlicher Gehälter von rund 1,40 USD pro

Stunde in Kenia (Perrigo 2023). Sama befasste sich im Auftrag von OpenAI mit schriftlichen sowie bildlichen Trainingsdaten zu den Themen sexueller Gewalt und Hassrede für KI Content Filter, und sammelte 1.400 Bilder, darunter Kindesmissbrauch und Bestialität, um KI „sicherer zu machen“, für den Preis, Clickworker dem Risiko eines Arbeitstrauma auszusetzen (Perrigo 2023). Es sei kein Einzelfall, dass Clickworker beim Moderieren von gewaltvollem Content psychische Schäden erfahren können (vgl. u.a. Siapera 2022).

Arbeitsaufträge der Clickworker

Zu den Aufträgen – sogenannte Mikro-Tasks – gehören Audio-Transkriptionen, Übersetzungen, Bildbeschreibungen, Kategorisierungen, Vergeben von Ratings, Korrigieren der Antworten von Assistent:innen, Analysieren von Facebookseiten, Ausfüllen von Fragebögen, Suchmaschinenoptimierungen, Promotionen – zumeist einfache, repetitive Arbeit (Grohmann und Fernandes Araujo 2021, 249–55; Tubaro, Le Ludec, und Casilli 2020, 67–68; Rani und Furrer 2021, 213; Hagendorff 2022, 555; Tubaro und Casilli 2019, 340; Miceli u. a. 2024, 4; Dzieza 2023). Clickworker schulen KI im Bereich des maschinellen Sehens, indem sie Aufgaben wie Bildklassifizierung, Objekterkennung, Markierung von Orientierungspunkten, Tagging und Segmentierung auf Pixel-ebene durchführen (Tubaro und Casilli 2019, 340). Diese Aufträge bestehen meistens darin, Verknüpfungen zwischen Material und Bedeutungen herzustellen, wie z.B. Bilder mit Attributen zu annotieren oder auszuwerten (Grohmann und Fernandes Araujo 2021, 251).

Diese Aufträge werden standardmäßig von drei bis 5 Clickworkern bearbeitet; im Falle unterschiedlicher Antworten werden mithilfe eines Mehrheitswahlsystems die Option mit der meisten Zustimmung als korrekt evaluiert (Rani und Furrer 2021, 222). Auch können Online-Plattformen die Produktion persönlicher Daten fordern, wie z.B. das Aufzeichnen der Stimme oder das Aufnehmen von Videos und Fotos des eigenen Körpers der Clickworker durch eine Kollektion an Selfies (vgl. Huws 2014; Grohmann und Fernandes Araujo 2021, 255; Miceli u. a. 2024, 4; Lulamae 2023). Prognosen zufolge werden diese Tätigkeiten einerseits in den kommenden Jahren weiterhin Bestand haben und andererseits voraussichtlich spezifischere und komplexere Herausforderungen umfassen, sobald eine bestimmte Annotationsarbeit obsolet wird (Tubaro und Casilli 2019, 341–42; Dzieza 2023).

Trotz der monotonen Arbeit der Clickworker widerspricht Dzieza im Kontext der Online-Plattformen dem Konzept der „Bullshit-Jobs“ von David Graeber: Im Gegensatz zu bedeutungslosen Jobs, die in Zukunft automatisiert werden sollten, bilden diese Mikro-Tasks den Grundbaustein für KI und benötigen stets menschliche Arbeit (Dzieza 2023).

Arbeitsverhältnisse der Clickworker

Für Online-Plattformen müssen Bewerber:innen in einem Selektionsprozess zunächst ihren Lebenslauf einreichen und werden durch unbezahlte Qualifizierungstests auf ihre Englisch-sprachigen Kompetenzen sowie die Projektrichtlinien geprüft (Grohmann und Fernandes Araujo 2021, 256; Rani und Furrer 2021, 217; Kretschmer 2024). Obwohl die meisten Aufträge der Online-Plattformen sowie ihre Qualifizierungshürden in Englischer Sprache sind, gibt es einige Clickworker, die keine Englischen Sprachkompetenzen beherrschen und stattdessen maschinelle Übersetzungen verwenden (Grohmann und Fernandes Araujo 2021, 260).

Vorteile der Arbeit auf Online-Plattformen seien der einfache, digitale Zugang und die zeitliche Flexibilität der Clickworker (Grohmann und Fernandes Araujo 2021, 255; Rani und Furrer 2021, 217). Da die meisten Online-Plattformen den Vereinigten Staaten angehören, wird der Großteil der Mikro-Tasks allerdings auch zu deren Tageszeiten veröffentlicht, weshalb Clickworker auf der anderen Seite des Globus sich zeitlich dem Workload sowie Konkurrenzdruck anpassen müssen (Rani und Furrer 2021, 225). Nichtsdestotrotz bestehen viele Arbeitnehmer:innen in Latein Amerika auf die von Online-Plattformen versprochene Autonomie, Flexibilität sowie das Selbstmanagement und lehnen bewusst eine Festanstellung ab (Grohmann und Fernandes Araujo 2021, 251).

Nachteile des Arbeitsverhältnisses seien die Absenz jeglicher Standards, der Mangel an Sicherheit bei der Einstellung, die Termination oder Nichtbezahlung der Arbeit aufgrund von falsch-klassifizierten Antworten (Grohmann und Fernandes Araujo 2021, 256; Tan und Cabato 2023). Dadurch, dass die Clickworker-Industrie ein relativ neues Arbeitsfeld bildet und zum informellen Sektor gehört, gibt es noch kaum regulative Schutzmechanismen für die Arbeitnehmer:innen (vgl. Tan und Cabato 2023). Auf Online-Plattformen sei es keine Seltenheit, dass Clickworker falschen Anschuldigungen oder bestimmter Dysfunktionalitäten der App ausgesetzt sind, und ihre Löhne nicht (voll) oder verzögert ausgezahlt werden (Grohmann und Fernandes Araujo 2021, 256–57; Tan und Cabato 2023). Beschwerden können zu einer Deaktivierung des Accounts führen (Tan und Cabato 2023; Dzieza 2023). Abhängig von der Region werden Löhne meist nur unter dem Mindestlohn bezahlt (Tan und Cabato 2023). Das Oxford Internet Institut z.B. beschrieb ScaleAI ihre Arbeitsprozesse zu verschleiern und bewertete Remotasks mit einem Score von 1 von 10 (Tan und Cabato 2023). Aufgrund des Mangels an Transparenz besteht eine erhöhte Schwierigkeit darin festzustellen, was eigentlich zur Arbeitszeit zählt (Rani und Furrer 2021, 217). Nach Schätzungen verbringen Clickworker rund acht unbezahlte Stunden jede Woche unter anderem für das Suchen neuer Aufträge (Kretschmer 2024).

In allem existiert ein großes Machtungleichgewicht zwischen Arbeitgeber:innen und Auftragnehmer:innen (Kretschmer 2024). Clickworker unterliegen Geheimhaltungsvereinbarungen, welche ihre Redefreiheit einschränken (Miceli u. a. 2024, 4). Clickworker werden zum Schutz vertraulicher Informationen konstant digital überwacht (Miceli u. a. 2024, 4). Bei Cloudfactory müssen sie für ihre Arbeit z.B. eine bestimmte Version des Chrome-Browsers von Google benutzen, wodurch Online-Plattformen Zugriff auf den Bildschirm sowie die Webcam haben; dies bestätigen die Arbeitsverträge (Kretschmer 2024).

Profil der Clickworker

Nach Schätzungen gibt es zur Zeit etwa 10 Millionen Clickworker weltweit, Tendenz steigend (Kretschmer 2024). Um die monotone Arbeit von Clickworkern zu beschreiben, gibt es folgende Synonyme und Metaphern, die in der Literatur verwendet werden: Clickworker oder Micro-Worker als „unsichtbare Arbeiter:innen“ absolvieren „Crowd-work“ oder „Ghost-work“ oft in Internetcafés betitelt als „digital sweatshops“ und bilden das „back-office“ für die Entwicklung von KI (Tubaro und Casilli 2019, 334; Rani und Furrer 2021, 217; Gray und Suri 2019; Grohmann und Fernandes Araujo 2021, 249; Tan und Cabato 2023). Clickworker sind unabhängige, anonymisierte und deshalb leicht auswechselbare, meistens niedrig-qualifizierte, niedrig-bezahlte Auftragnehmer:innen häufig im Homeoffice, die von Online-Plattformen rekrutiert werden, um im Auftrag von KI-Technologiekonzernen Mikro-Tasks zu bearbeiten (Tubaro und Casilli 2019, 334; Rani und Furrer 2021, 217). Diese Anonymität zwischen Arbeitnehmer:innen und Arbeitgeber:innen vereinfacht eine algorithmische Verwaltung menschlicher Arbeit und verschleiert Ausbeutung (Rani und Furrer 2021, 217; Grohmann und Fernandes Araujo 2021, 248).

Dabei gibt es sowohl experimentierfreudige Individuen, die die Arbeit auf Online-Plattformen lediglich ausprobieren, als auch regelmäßige sowie ständig aktive Clickworker, die Mikro-Tasks als Fulltime-Job ablegen (Tubaro, Le Ludec, und Casilli 2020, 72). Frauen mit Kindern oder Arbeitslosen haben eine Möglichkeit von zuhause aus einen Minijob anzunehmen (Tubaro und Casilli 2019, 4; Al-Sibai 2023). Häufig bedienen Online-Plattformen jedoch einen Arbeitsmarkt mit einer hohen Arbeitslosenquote sowie niedrigen Arbeitslöhnen in Ländern des globalen Südens, aber auch in Europa (Miceli u. a. 2024, 5). Die Ausländerrate auf Online-Plattformen in Europa ist mit 39% relativ hoch, vermutlich weil viele Migrant:innen trotz hohem Bildungsgrad keine Arbeit finden (vgl. Miceli u. a. 2024, 5). Obwohl die Arbeitsverhältnisse auf internationaler Ebene nicht homogen sind (Grohmann und Fernandes Araujo 2021, 251), erweckt die Clickworker-Branche den Eindruck eine neue Arbeiterklasse hervorgebracht zu haben (Kretschmer 2024).

Viele Clickworker sehen sich nicht als Teil der KI-Industrie (Grohmann und Fernandes Araujo 2021, 261; Lulamae 2023) und stattdessen als Opfer der Ausbeutung oder Sklaverei (Grohmann und Fernandes Araujo 2021, 253; Al-Sibai 2023). Darüber hinaus sind Clickworker einem erhöhten Risiko ausgesetzt, infolge des Arbeitsdrucks ohnmächtig zu werden, einen Burnout zu erleiden, psychotische Episoden zu erleben, unter Isolation zu leiden und Suizid zu begehen (Miceli u. a. 2024, 5). Nur ein geringer Anteil dieser Arbeiter:innen besitzt eine Gesundheits- oder Sozialversicherung (Rani und Furrer 2021, 230). Obwohl das EU-Datenschutzgesetz DSGVO existiert, ist dessen Umsetzung in einem intransparenten Arbeitsverhältnis schwierig, zumal Auftragsnehmer:innen sich ihrer Rechte meist unbewusst sind (Lulamae 2023).

KI-Entwicklung entlang der globalen Asymmetrie

Sprache als Mittel zur Erschließung von Arbeitsmärkten

Dieses Unterkapitel beschäftigt sich mit Online-Plattformen, die entlang der Nord-Süd-Asymmetrie des Wohlstands anhand der Sprache als Mittel sich Zugang zu Arbeitsmärkten in Ländern des globalen Südens verschaffen. Die westliche Entwicklung von KI wird häufig dafür kritisiert, die koloniale Infrastruktur zu nutzen und koloniale Praktiken zu reproduzieren. Westliche Online-Plattformen beziehen ihre Trainingsdaten überwiegend aus Ländern des globalen Südens, die aufgrund historisch bedingter imperialer Eingriffe mit der englischen Sprache vertraut gemacht wurden und einen günstigen Arbeitsmarkt bieten. Die Tatsache, dass große Teile der Bevölkerung in den ehemals kolonialisierten Ländern aufgrund dieser historischen Einflüsse entweder natürlich Englisch sprechen oder Englisch als Verkehrssprache erlernen, erweist sich dabei als zufällig vorteilhafte Fügung, da Englisch als Zugang zum Arbeitsmarkt genutzt werden kann. Die englische Sprache ist unter anderem eine essenzielle Voraussetzung für die Produktion von kompatiblen Trainingsdaten für sowohl Text- als auch Bildgeneratoren, weil beide auf Annotationsarbeiten angewiesen sind.

Als koloniale Praxis wird unter anderem die sogenannte „algokratische“ Arbeitsstruktur (algocratic governance) verstanden; diese beschreibt im Kontext der Online-Plattformen eine Art, eine Masse von Clickworkern auf effiziente Weise durch automatisierte Prozesse und Rating- sowie Reputationssysteme algorithmisch zu organisieren (Rani und Furrer 2021, 213–16). Diese algorithmische Form des Managements gestaltet prekäre Arbeitsverhältnisse und Konditionen für Clickworker (Rani und Furrer 2021, 231).

Im Kontext der Digitalisierung, Sozialer Medien und der Content-Moderation in Ländern des globalen Südens durch Online-Plattformen verweisen Udupa und Datatreyan auf einen Überwachungskapitalismus, in dem personenbezogene Daten

kommodifiziert und kommerzialisiert werden (Udupa und Dattatreya 2023, 97 ff.). Sie führt das Konzept des „Capture“ ein, das die internationale Extraktion und Aneignung persönlicher Daten als proprietären Verhaltensüberschuss („proprietary behavioral surplus“) beschreibt, was an koloniale Praktiken erinnert (Udupa und Dattatreya 2023, 99). Dass asymmetrische Datenbeziehungen zwischen Nationalstaaten vorherrschen und sich somit ein Machtgefälle etabliert, stößt in der Literatur auf einen allgemeinen Konsens (vgl. Siapera 2022; Tubaro und Casilli 2019, 334).

Während westliche Social-Media-Plattformen im globalen Süden personenbezogene Daten extrahieren, um nicht zuletzt KI zu füttern, bilden Online-Plattformen dafür ein digitales Instrument, billige Arbeit zu extrahieren (Grohmann und Fernandes Araujo 2021, 260 ff.). In diesem Sinne sind Online-Plattformen ein effizienter Mechanismus den Surplus-Wert anspruchsloser Click-Arbeit zu maximieren und im Anbetracht der lukrativen KI-Industrie hohe Rendite einzufahren. Online-Plattformen wählen darüber hinaus bewusst Arbeitsmärkte in Ländern des globalen Südens aus, weil dort die Sozialsysteme noch relativ schwach sind und der Druck zur Bereitstellung von sozialen Leistungen gering ist (Rani und Furrer 2021, 230).

Die Kombination aus der staatlich-subventionierten Digitalisierung wie z.B. in Indien, Brasilien und auf den Philippinen, ihre Integration in die Weltwirtschaft mit geringen politischen Reglementierungen und ihre wohlwollende Aneignung englischer Kompetenzen, schafft attraktive Pull-Faktoren für das Outsourcen westlicher Billigarbeit, die meist langfristig nicht im Interesse der Gesamtbevölkerung sind (Rani und Furrer 2021, 215). Philippinische Arbeitsgruppen beschwerten sich bei ihrer Regierung über das Defizit einer Regulierung der Online-Plattformen, doch Beamt:innen fürchten die schnell-wachsende Branche zu lähmen (Tan und Cabato 2023). Dass Microwork als „dead-end jobs“ betrachtet werden, angesichts des Mangels nach Karrieremöglichkeiten sowie übertragbaren Fähigkeiten für andere Berufe, wird in vielen wirtschaftsschwachen Ländern als kleineres Übel geduldet, im Hinblick auf die ohnehin hohe Arbeitslosenquote (Rani und Furrer 2021, 214–16). Es ist davon auszugehen, dass das Outsourcen an Mikro-Tasks langfristig zu einer Dequalifizierung (deskilling) von Arbeiter:innen führt (Rani und Furrer 2021, 217).

Um die gravierenden ökonomischen und sozialen Auswirkungen bestimmter sprachlicher Regionen zu analysieren, kann beobachtet werden, inwieweit nicht-englischsprachige Nationalstaaten KI-Systeme in ihrer jeweiligen Sprache trainieren. Französische Unternehmen betreiben teilweise Outsourcing von Klickarbeit entlang kolonialer Verbindungen in ehemalige Kolonien auf dem afrikanischen Kontinent (Tubaro und Casilli 2019, 343). Finnische Technologieunternehmen hingegen beginnen aufgrund der geringen Verbreitung ihrer Sprache, kostengünstige Arbeitskräfte in heimischen Gefängnissen für die Produktion von KI-Trainingsdaten einzusetzen; Gefangene

dürfen während ihres unfreiwilligen Aufenthalts Daten für 1,54 Euro pro Stunde annotieren (Taylor 2023). In Chinas KI-Industrie ist aufgrund der sprachlichen Besonderheiten das Outsourcing von Mikroaufgaben in andere Länder ebenfalls kaum möglich. Chinesische Technologieunternehmen verpflichten Berufsschüler:innen im Rahmen ihrer Ausbildung dazu, KI-Trainingsdaten unter ausbeuterischen Löhnen zu annotieren (Zhou und Chen 2023).

Vertikale & Horizontale Arbeitsteilung

In diesem Unterkapitel geht es um (1) die vertikale, internationale Arbeitsteilung zwischen komplexer sowie einfacherer Arbeit entlang der Nord-Süd-Asymmetrie und (2) die Auswirkungen der Online-Plattformen bei der horizontalen Verteilung von Aufträgen durch die Etablierung eines globalen Arbeitsmarktes unter direkter, internationaler Konkurrenz, entlang der Süd-Achse.

(1) Nach Angaben des Online Labour Index (OLI) der Universität von Oxford seien die USA das Land mit den meisten Online-Aufträgen und Indien das Land mit den meisten Online-Auftragnehmer:innen (Grohmann und Fernandes Araujo 2021, 249). Erik Duhaime, CEO des Unternehmens Centaur Lab, erklärt: „AI doesn't replace work [...] But it does change how work is organized.“ (Dzieza 2023). Dies bezieht sich nicht nur auf die Möglichkeiten der Verteilung von Mikro-Tasks an Crowdworkern durch Online-Plattformen oder des algorithmischen Managements, sondern auch generell auf die internationale Arbeitsteilung entlang einer Nord-Süd-Achse. In der wissenschaftlichen Literatur wird thematisiert, dass die Entwicklung von KI von einer eher weißen, männlichen Elite wie z.B. im Silicon Valley geprägt wird, wobei der Grad der Mitbestimmung kontinuierlich abnimmt, je weiter man sich südlich bewegt (vgl. Udupa und Dattatreyan 2023, 6; Siapera 2022).

Viele Online-Plattformen beschränken z.B. den Zugang zu bestimmten Aufträgen, sodass Clickworker im Westen tendenziell besser bezahlte Aufträge erhalten, während Clickworker in Ländern des globalen Südens häufiger gesperrt werden und mehr Zeit benötigen, um eine ausreichende Reputation zu erlangen, die den Zugang zu besser bezahlten Jobs ermöglicht (Rani und Furrer 2021, 216–17). Dies führt zu einem erheblichen Einkommensunterschied zwischen den Regionen auf derselben Plattform und verursacht geographische Diskriminierung (Rani und Furrer 2021, 222–24). Nach Marx führe eine derartige Arbeitsteilung insofern zu ihrer Entfremdung, dass der Beitrag der Clickworker an der Entwicklung der KI unnachvollziehbar wird (Grohmann und Fernandes Araujo 2021, 251). Im Hinblick auf den marxschen Warenfetischismus verschleiern KI-Konzerne ihre Beziehung zu den Clickworkern über eine lange, anonymisierte Lieferkette.

Die Tatsache, dass Arbeitskräfte in Ländern des globalen Südens tendenziell günstiger sind als im Westen, erscheint mir logisch in der Wirtschaftstheorie, doch

auch insofern banal, dass Clickworker ein VPN benutzen können, um die geographische Diskriminierung zu umgehen und besser-bezahlte Aufträge zu erhalten. Erkennt eine Online-Plattform dieses Vorgehen bei einem Clickworker, wird der Account sofort gesperrt (Dzieza 2023).

(2) Entlang einer Süd-Achse sorgt die direkte Konkurrenz internationaler Clickworker auf einem globalen Arbeitsmarkt für ein erhöhtes Angebot an verfügbaren Arbeitskräften im Verhältnis zu wenigen Aufträgen, wodurch der Lohn pro Auftrag sinkt. Nachdem z.B. die Online-Plattform Remotasks ihre Auftragsvermittlung auf Indien und Venezuela ausweitete, nahmen die durchschnittlichen Löhne pro Auftrag von 10 USD auf 1 Cent ab (Tan und Cabato 2023). Die Möglichkeit auf diese Weise einen globalen Arbeitsmarkt zu etablieren, beruht nicht zuletzt auf der historisch-bedingten Verbreitung der englischen Kommunikationsprache. Angesichts dieser universellen Verbreitung ist Englisch eine günstige Sprache, mithilfe dessen westliche Länder einen ungesättigten Arbeitsmarkt an billigen Clickworkern für die KI-Entwicklung rekrutieren können. So bezog z.B. ChatGPT-3 zu 93% Trainingsdaten aus der englischen Sprache (Brown u. a. 2020).

Zynische Ironie der Produktion von Trainingsdaten

Die zynische Ironie der Produktion von Trainingsdaten liegt einerseits darin, dass etwa zwischen 33% und 46% der Clickworker für ihre Mikro-Tasks ChatGPT verwenden, Tendenz steigend, wodurch KI in einer Schleife andere KI füttern und das Risiko von sich amplifizierenden Fehlern erhöht (Williams 2023). Andererseits liegt eine böse Ironie darin, dass Clickworker in Ländern des globalen Südens westliche KI trainieren, die vermutlich in Zukunft wiederum ihr Leben nachhaltig prägen oder sogar einschränken können. Nach Hagendorff sei KI ein Herrschaftsinstrument oft in den Händen von mächtigen Personengruppen zur Kontrolle von Menschen (Hagendorff 2022, 852).

Ein Beispiel sei, wie ausländische Clickworker mit Aufnahmen ihrer Stimme eine Erkennungssoftware für Dialekte trainieren, über Online-Plattformen im Auftrag des deutschen Bundesamtes für Migration und Flüchtlinge (BAMF) (Lulamae 2022; 2023). Diese Erkennungssoftware soll in Zukunft Behörden helfen, die Identität und Herkunft von Flüchtlingen zu verifizieren, und werde Sprachen wie Farsi, Dari, Paschtu, Irakisches Arabisch, Maghrebi-Arabisch, Levantinisches Arabisch, Golf-Arabisch und Ägyptisches Arabisch unterscheiden können, obwohl eine solche Identifikation als Pseudowissenschaft gilt (Lulamae 2022). Dies führt zu einer banal-rassistischen Pervertierung, die zum Denken anregt, inwiefern Herkunft und Identität so rational kategorisiert werden muss. Letzten Endes trainieren ausländische Clickworker eine fehleranfällige KI, die in den Augen von deutschen Behörden objektiv-richtige Angaben über Eigenschaften fremder Personen produzieren und eventuell in Zukunft

darüber entscheiden, ob genau diese Clickworker oder dessen ethnische Gruppe berechtigt sind, nach Deutschland (oder Europa) einzureisen.

Ein anderes Beispiel sei die Kollaboration von ScaleAI mit dem Verteidigungsministerium der USA (Tan und Cabato 2023; Dzieza 2023). Im Anbetracht der Effektivität von Kampfdrohnen in Kriegsgebieten wie der Ukraine und Russland, würde sich der zynisch-ironische Kreis schließen, wenn US-Unternehmen Clickworker in Ländern des globalen Südens zu beauftragen, Trainingsdaten für autonome Drohnen zu produzieren, die potentiell wiederum einer militärischen Agenda dienen könnten, womit Clickworker niemals zustimmen konnten.

Schluss

Im Kontext des wettbewerbsorientierten Laufs um die Entwicklung von KI und dem damit verbundenen Bedarf an Daten untersucht diese Arbeit den Prozess der Produktion von Trainingsdaten durch Clickworker. Der wissenschaftliche Beitrag dieser Untersuchung liegt darin, die historisch bedingte Verbreitung der englischen Sprache als einen zentralen Faktor für die Etablierung und Manifestierung sozial-struktureller Arbeitsverhältnisse zwischen Ländern des globalen Südens und Nordens hervorzuheben. Online-Plattformen spielen eine entscheidende Rolle bei der Strukturierung von Arbeitsmärkten, indem sie Mikroaufträge verwalten und vermitteln. Dabei diskriminieren sie auf einer vertikalen Achse geografisch ökonomisch schwache Bevölkerungsgruppen und senken auf einer horizontalen Achse die Lohnkosten durch direkte internationale Konkurrenz. Selbst wenn man historisch bedingte Ursachen wie koloniale und imperiale Eingriffe westlicher Länder sowie das daraus resultierende Ungleichgewicht im erlangten Wohlstand außer Acht lässt und langfristige Abhängigkeitsverhältnisse wie unausgeglichene Datenbeziehungen ignoriert, bestehen weiterhin gravierende Missstände. Dazu zählen mangelnde Transparenz und unbezahlte Arbeit, die von Online-Plattformen als unvermeidliche Kollateralschäden hingenommen werden und aufgrund des Machtgefälles zwischen Auftraggeber:innen und Auftragnehmer:innen kaum rechtlich angefochten werden können. Als Anthropologe ist es besonders frustrierend zu beobachten, wie die KI-Industrie durch Online-Plattformen erneut auf die altbekannte koloniale Infrastruktur zurückgreift, um wirtschaftliche Vorteile aus den Ländern des globalen Südens zu ziehen, bis die betroffenen Subalternen ihre oft korrupten Regierungen davon überzeugen können, diese Tendenzen zu regulieren.

Literatur

- Al-Sibai, Noor. 2023. „That AI You’re Using Was Trained By Slave Labor, Basically“. *Futurism*. 20. Oktober 2023. <https://futurism.com/the-byte/ai-gig-slave-labor>.
- Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, u. a. 2020. „Language Models are Few-Shot Learners“. *arXiv*. <https://doi.org/10.48550/arXiv.2005.14165>.
- Dzieza, Josh. 2023. „Inside the AI Factory“. *The Verge*. 20. Juni 2023. <https://www.theverge.com/features/23764584/ai-artificial-intelligence-data-notation-labor-scale-surge-remotasks-openai-chatbots>.
- Graham, Mark, Isis Hjorth, und Vili Lehdonvirta. 2017. „Digital Labour and Development: Impacts of Global Digital Labour Platforms and the Gig Economy on Worker Livelihoods“. *Transfer: European Review of Labour and Research* 23 (2): 135–62. <https://doi.org/10.1177/1024258916687250>.
- Grand View Research. 2023. „Data Collection And Labeling Market Worth \$17.10 Billion By 2030“. Grand View Research. März 2023. <https://www.grandviewresearch.com/press-release/global-data-collection-labeling-market>.
- Gray, Mary L., und Siddharth Suri. 2019. *Ghost Work : How to Stop Silicon Valley from Building a New Global Underclass*. Boston : Houghton Mifflin Harcourt.
- Grohmann, Rafael, und Willian Fernandes Araujo. 2021. „Beyond Mechanical Turk: The Work of Brazilians on Global AI Platforms“. In , 247–66. <https://doi.org/10.16997/book55.n>.
- Hagendorff, Thilo. 2022. „Blind Spots in AI Ethics“. *AI and Ethics* 2 (4): 851–67. <https://doi.org/10.1007/s43681-021-00122-8>.
- Hornborg, Alf. 2014. „Technology as Fetish: Marx, Latour, and the Cultural Foundations of Capitalism“. *Theory, Culture & Society* 31 (4): 119–40. <https://doi.org/10.1177/0263276413488960>.
- Hornuf, Lars, und Daniel Vrankar. 2022. „Hourly Wages in Crowdsourcing: A Meta-Analysis“. *Business & Information Systems Engineering* 64 (5): 553–73. <https://doi.org/10.1007/s12599-022-00769-5>.
- Huws, Ursula. 2014. „Monthly Review | Labor in the Global Digital Economy: The Cybertariat Comes of Age“. *Monthly Review* (blog). 4. Dezember 2014. https://monthlyreview.org/product/labor_in_the_global_digital_economy/.
- Javaid, Shehmir. 2024. „Top 12 Data Crowdsourcing Platform: Evaluation & User Reviews“. *Research AIMultiple*. 3. Januar 2024. <https://research.aimultiple.com/data-crowdsourcing-platform/>.
- Kretschmer, Christian. 2024. „Clickworker – Ausgebeutet für künstliche Intelligenz“. *swr.online*. 3. Juli 2024. <https://www.swr.de/swrkultur/wissen/clickworker-ausgebeutet-fuer-kuenstliche-intelligenz-104.html>.
- Lulamae, Josephine. 2022. „The BAMF’s Controversial Dialect Recognition Software: New Languages and an EU Pilot Project“. *AlgorithmWatch*. 5. September 2022. <https://algorithmwatch.org/en/bamf-dialect-recognition/>.
- — —. 2023. „A Dollar for Your Face: Meet the People behind Machine Learning Models“. *AlgorithmWatch*. 23. Februar 2023. <https://algorithmwatch.org/en/people-behind-machine-learning-models/>.
- Miceli, Milagros, Paola Tubaro, Antonio A. Casilli, Thomas Le Bonniec, Camilla Salim Wagner, und Laurenz Sachembacher. 2024. „Who Trains the Data for European Artificial Intelligence?“ *European Parliament ; The Left*. <https://hal.science/hal-04662589>.

- Perrigo, Billy. 2023. „Exclusive: The \$2 Per Hour Workers Who Made ChatGPT Safer“. TIME. 18. Januar 2023. <https://time.com/6247678/openai-chatgpt-kenya-workers/>.
- Rani, Uma, und Marianne Furrer. 2021. „Digital Labour Platforms and New Forms of Flexible Work in Developing Countries: Algorithmic Management of Work and Workers“. *Competition & Change* 25 (2): 212–36. <https://doi.org/10.1177/1024529420905187>.
- Siapera, Eugenia. 2022. „AI Content Moderation, Racism and (de)Coloniality“. *International Journal of Bullying Prevention* 4 (1): 55–65. <https://doi.org/10.1007/s42380-021-00105-7>.
- Tan, Rebecca, und Regine Cabato. 2023. „Behind the AI Boom, an Army of Overseas Workers in ‘Digital Sweatshops’“. *Washington Post*, 28. August 2023. <https://www.washingtonpost.com/world/2023/08/28/scale-ai-remotasks-philippines-artificial-intelligence/>.
- Taylor, Ben. 2023. „Analysis: Long Hours and Low Wages: The Human Labour Powering AI’s Development“. *Brighter World* (blog). 16. November 2023. <https://brighter-world.mcmaster.ca/articles/analysis-long-hours-and-low-wages-the-human-labour-powering-ais-development/>.
- Tubaro, Paola, und Antonio A. Casilli. 2019. „Micro-Work, Artificial Intelligence and the Automotive Industry“. *Journal of Industrial and Business Economics* 46 (3): 333–45. <https://doi.org/10.1007/s40812-019-00121-1>.
- . 2020. „Portraits of micro-workers: The real people behind AI in France“. In *2nd Crowdfunding Symposium 2020*. Paderborn, Germany: Research program “Digital Future”, Universities of Paderborn and Bielefeld. <https://hal.science/hal-02960775>.
- Tubaro, Paola, Clément Le Ludec, und Antonio A. Casilli. 2020. „Counting ‘micro-workers’: societal and methodological challenges around new forms of labour“. *Work Organisation, Labour & Globalisation* 14 (1): 67–82. <https://doi.org/10.13169/workorgalaboglob.14.1.0067>.
- Udupa, Sahana, und Ethiraj Gabriel Dattatreyan. 2023. *Digital Unsettling: Decoloniality and Dispossession in the Age of Social Media*. 1. Aufl. New York: University Press.
- Williams, Rhiannon. 2023. „The People Paid to Train AI Are Outsourcing Their Work... to AI“. MIT Technology Review. 22. Juni 2023. <https://www.technologyreview.com/2023/06/22/1075405/the-people-paid-to-train-ai-are-outsourcing-their-work-to-ai/>.
- Zhou, Viola, und Caiwei Chen. 2023. „China’s AI Boom Depends on an Army of Exploited Student Interns“. Rest of World. 14. September 2023. <https://restofworld.org/2023/china-ai-student-labor/>.