«Wer dem KI-Hype verfällt, stärkt die Macht der Big-Tech-Chefs»

Künstliche Intelligenz werde massiv überhöht, sagt Signal-Präsidentin Meredith Whittaker. Ein Gespräch über die Agenda der Silicon-Valley-Konzerne und die gefährlichen Pläne der EU.

Von Adrienne Fichter (Text) und Caitlin Chescoe (Bilder), 05.07.2023

«Die Macht- und Besitzfrage ist die eigentliche Frage, die gestellt werden muss»: Meredith Whittaker.

Frau Whittaker, zurzeit erscheinen gefühlt im Monatstakt offene Briefe, die vor der Vernichtung der Menschheit durch künstliche Intelligenz, kurz KI, warnen. Die Absender sind Big-Tech-Milliardäre wie Elon Musk oder auch Chat-GPT-Gründer Samuel Altman. Was steckt dahinter?
Der Hype rund um künstliche Intelligenz hat schon religiöse Ausmasse angenommen. Es gibt keine Evidenz dafür, dass KI-Technologien jemals ein Bewusstsein erlangen oder superintelligent sein werden. Was uns Sorge machen sollte, sind die riesigen Infrastrukturen, die im Besitz dieser Warner sind.

Wie wurde das möglich, dass wenige Tech-Unternehmen solche massiven Infrastrukturen bauen konnten?
Diese Entwicklung begann bereits in den 1990er-Jahren. Damals beauftragte die Clinton-Regierung Expertinnen mit der Bewertung von Risiken und Chancen der vernetzten Dateninfrastruktur. Doch statt auf die eigenen Experten zu hören, verfiel die Clinton-Administration dem neoliberalen Glauben, dass man die Besitzer dieser Technologien einfach gewähren lassen soll. Hightech sei Balsam für eine kränkelnde Wirtschaft. Und dadurch konnte sich auch der Überwachungskapitalismus ungehindert entfalten. Heute gibt es also einige wenige Big-Tech-Firmen mit riesigen Server-Infrastrukturen, enorm hohen Datenspeicherkapazitäten und vielen Milliarden Nutzerinnendaten, mit denen dann simpel gesagt statistische Model-

le erstellt werden. Die Macht- und Besitzfrage ist die eigentliche Frage, die gestellt werden muss.

Sie sagten in einer Rede an der Digitalkonferenz «re:publica» in Berlin: Wer an den KI-Hype glaube, mache die Mächtigen noch mächtiger. Allein schon die wiederholte Erzählung dieses Mythos entfalte seine Wirkung.
Wir dürfen ja alle ruhig an etwas glauben, für das es keine Beweise gibt. Doch wir müssen uns bewusst sein, dass es Akteure gibt, die enorm von diesen Narrativen profitieren. Wer also diesem KI-Hype verfällt, stärkt die infrastrukturelle Macht der Big-Tech-Chefs. Nochmals: Wir reden hier über geballte unternehmerische und wirtschaftliche Power. Es geht um die Nutzung grosser Datenmengen und um die Entwicklung statistischer Modelle, die Entscheidungen über uns und unsere Welt treffen sollen. Und dies soll ohne jede Rechenschaftspflicht gegenüber der Bevölkerung geschehen. Zum Nutzen dieser Profiteure – und nicht im Dienste des Gemeinwohls. Das ist die wirkliche Gefahr, die wir im Auge behalten müssen. Für den Rest gibt es keine Evidenz.

Zur Person

Meredith Whittaker ist eine renommierte Expertin für ethische Fragen rund um den Datenschutz und künstliche Intelligenz. Sie ist seit vergangenem Jahr Präsidentin der Stiftung Signal, die die gleichnamige populäre Messenger-App anbietet. Davor arbeitete sie von 2006 bis 2019 bei Google, wo sie die Google Open Research Group gründete. Sie war treibende Kraft hinter den Protesten von Google-Mitarbeiterinnen gegen Diskriminierung am Arbeitsplatz, Überwachung und militärische Projekte des Konzerns.

Auch wenn die Erzählung der «die Menschheit auslöschenden KI-Systeme» Quatsch sein sollte: Immerhin fordert das Silicon Valley eine Regulierung und Regeln für den Einsatz dieser KI-Systeme. Das ist doch eine gute Sache.
Regulierung ist ein wirklich grosser Begriff. Es gibt auch schlechte Regulierung, zum Beispiel die «Notice and Consent»-Richtlinie in den USA mit all den Cookie-Bannern, die wir jetzt abnicken müssen, um überhaupt auf eine Website zu gelangen. Damit wird die ungewollte Datenverarbeitung aber nicht verhindert. Wir haben es bereits bei der Tabakindustrie gesehen: Wenn klar wird, dass Regulierung unausweichlich ist, bringen sich die Unternehmen selber massiv ein. Denn die beste Alternative zu «keiner Regulierung» für sie ist: die Regulierung selber mitzuschreiben. Insofern sind die offenen Briefe des Silicon Valley eine Art riesige Influencer-Kampagne.

Was ist die Strategie dahinter?
All die Unternehmenschefs, die Lobbyisten, teils auch die befreundeten, von Big Tech finanzierten Akademikerinnen, welche jetzt vor Menschheitskatastrophen warnen, möchten eigentlich keine Beschränkung der heutigen Technologien. Sie wollen, dass der Status quo erhalten bleibt. Und deshalb werden irgendwelche existenzvernichtenden KI-Systeme der Zukunft herbeifantasiert, damit wir uns nicht mit dem Jetzt beschäftigen. Denn die heutigen KI-Systeme sind bereits schädigend. Sie reproduzieren Stereotype, sie haben einen rassistischen und sexistischen Bias, weil sie ja mit Texten, Bildern und Videos aus dem gesamten Internet trainiert und gefüttert werden.

Also scheint es eine Strategie der Ablenkung von den echten heutigen Gefahren zu sein. Ist der geplante «AI Act» der EU, also das geplante KI-Gesetz, in Gefahr, wenn so viele einflussreiche Lobbyorganisationen von Microsoft und Google in Brüssel vor Ort mitreden?
Beim «AI Act» der EU werden KI-Technologien nach Anwendungsfeldern und Risikoszenarien kategorisiert und je nachdem strenger oder weniger streng reguliert. Ich bin kein grosser Fan dieses risikobasierten Ansatzes. Aus zwei Gründen: Er setzt sich zum einen nicht mit der Machtkonzentration bei den Betreibern und Besitzerinnen von KI-Systemen auseinander. Und zum anderen wäre es eine riesige verpasste Chance, wenn die EU die neuen, vermeintlich harmlosen Allzweck-KI-Modelle wie zum Beispiel Chat GPT nicht in dieser Hochrisikokategorie einstufen würde. Damit hätten Microsoft, Open AI und Google gewonnen. Denn je nach Verwendungszweck richten solche KI-Tools genauso Schaden an.

Doch sie könnten auch einen Nutzen haben. Elon Musk möchte mit Gehirnimplantaten Krankheiten ausmerzen. Sie kritisieren dies als eugenische Vision: Ein Milliardär bestimme, was als normal und gesund gelten solle. Doch wäre es nicht hilfreich, Technologien zur Emotionserkennung zu fördern zum Beispiel für Menschen mit Behinderungen, die nicht mehr sprechen können?
Lassen Sie uns zuerst darüber reden, wie Firmen solche Emotionserkennungssysteme erstellen: Sie lassen zuerst alle Bilder von Facebook als Datensatz herunterladen und stellen schliesslich einen Haufen prekär bezahlter Angestellter ein, die diese Bilder verschlagworten. Die Angestellten müssen das Bild jeweils acht Emotionen zuordnen. Ihre Aufgabe ist es also, zu bestimmen, welche Emotion das jeweilige Gesicht auf dem Foto ausdrückt. Ich trainiere damit ein KI-Modell aufgrund von Zuschreibungen, die diese Angestellten den Bildern geben. Angenommen, ich wäre jetzt eine solche Angestellte und schaue Sie an. Welche der acht Emotionen passt zu Ihnen? *[Republik-Reporterin schweigt und schaut leicht amüsiert.]* Oh, wütend. Sie sehen wütend aus. Oder Sie schauen gerade etwas skeptisch. Okay, habe ich Ihre Gefühle richtig gelesen?

Nein, ich war gerade leicht amüsiert …
Eben. Habe ich jetzt das Recht, Ihnen und der ganzen Welt zu sagen, wie Sie sich fühlen? Nein. Das Ganze wird noch absurder, wenn jemand eine Behinderung hat. Eine autistische Person zum Beispiel drückt sich mit einer anderen Mimik aus. Wie wird ein Angestellter die Emotionen dieser Person richtig erkennen?

Was wären weitere Grenzen der Emotionserkennung?
Nehmen Sie nur schon die kulturelle Varianz! Je nach Kultur, in der wir leben, und je nachdem, wie wir Emotionen verstehen, können wir sie ganz unterschiedlich ausdrücken. Ich komme aus den USA. Wir lachen gerne laut und sind damit manchmal ziemlich nervig. Jemand in Japan zeigt Freude ganz anders. Wenn wir also anerkennen, wie komplex, vielseitig und grossartig dieses Spektrum der Erfahrungen und des Bewusstseins ist, hilft uns das auch, zu erkennen, wie dumm eigentlich Systeme der künstlichen Intelligenz sind.

Unter dem Strich: Jene Technologien, die bei psychischen und physischen Krankheiten helfen sollen, richten mehr Schaden an?
Zumindest im kommerziellen Bereich haben solche Gefühlserkennungstechnologien überhaupt nichts mit Gedankenlesen zu tun, sondern sie basieren letztlich auf Annahmen einiger schlecht bezahlter Arbeiterinnen zu einem Gesicht einer fremden Person. Die Macher dieser Systeme werden Ihnen und den Unternehmen, die diese Systeme kaufen, aber weismachen: Sie wissen ja vielleicht nicht, dass Sie eigentlich wütend sind! Solche Erzäh-

lungen sind brandgefährlich. Es gibt null Evidenz dafür, dass es möglich ist, den inneren Zustand von Menschen irgendwie automatisiert zu lesen. Und wie gefährlich wird es, wenn diese dummen Systeme gravierende Auswirkungen auf das Leben der Menschen haben?

Sie erwähnen die menschliche Arbeit, die in der Entwicklung von Systemen wie Chat GPT steckt. Dabei geht es etwa um die Ausbeutung von Angestellten aus Kenia, die Texte über Gewalt, Suizid und Kindsmissbrauch für zwei Dollar pro Stunde aussortieren müssen. Was fordern Sie dabei konkret?
Mir geht es um Folgendes: Bei Textgeneratoren-Tools verbessert die von Menschen geleistete Arbeit – anders als bei Emotionserkennung – die gesamte Maschine. Wenn man der Maschine eine Intelligenz zuschreibt, verdrängt man die menschliche Intelligenz, die erforderlich war, um die nötigen Daten zu erstellen, sie richtig einzuordnen und das System so zu justieren, dass es weniger rassistische oder sexistische Resultate ausspuckt. Der vermeintliche Zauber rund um die KI-Systeme des Silicon Valley beruht also auf der Arbeit von menschlichem Personal, ohne das nichts von alledem funktionieren würde. Natürlich sollen diese Arbeiter mehr Geld bekommen, aber wenn wir das alles auf einen existenzsichernden Lohn hochrechnen ...

... dann wären diese KI-Tools schnell mal nicht mehr frei verfügbar und gratis.
Genau. Es gäbe keinen Markt dafür. Dann wären wir mit der echten ökonomischen Realität und einer Vollkostenrechnung konfrontiert. Solche Systeme würden sich schnell nicht mehr rentieren. Würden alle für den echten Wert ihrer Arbeit fair entschädigt werden, dann gäbe es kein Geschäftsmodell für diese Systeme. Und wir reden hier noch nicht einmal von den Künstlerinnen, die für den Klau ihrer Bilder zum Training der Modelle nicht entschädigt worden sind.

Kommen wir auf die Pläne der EU zu sprechen, die Ihre Arbeit beeinflussen werden. Messenger-Apps wie Signal, die Ende-zu-Ende-Verschlüsselung der Inhalte anbieten, sollen verpflichtet werden, die Inhalte ihrer Nutzerinnen zu durchsuchen und verdächtige Inhalte an ein EU-Zentrum zu senden.
Ich war natürlich negativ überrascht und dachte: echt jetzt? Ich komme aber aus den USA und konnte diese Meldung zuerst nicht einordnen im EU-Kontext.

13 EU-Staaten wollen verschlüsselte Inhalte durchleuchten. Dabei sollen alle eingehenden Nachrichten in unseren Messenger-Apps direkt auf unseren Smartphones auf Inhalte von sexuellem Kindsmissbrauch sowie Grooming (gezielte Kontaktaufnahme Erwachsener mit Minderjährigen) gescannt werden.
Da sind wir wieder beim Thema meiner Rede. Die Entscheidungsträgerinnen in Europa verfallen der KI-Magie. Das massenhafte Scannen unserer Smartphones ist ganz einfach eine Überwachungsmethode. Sie ist sehr stark fehleranfällig. Künstliche Intelligenz wird dieses Bild- und Textmaterial nicht erkennen. Politiker sind aber leider bereit, diesen Unsinn zu glauben. Auch hier spielen die Heilsversprechen rund um den Mythos KI wieder eine grosse Rolle. Das Thema Kindsmissbrauch ist ein mächtiger politischer Trigger, evidenzbasierte Entlarvung dieser Technologiemythen kommt dagegen nicht an.

«Signal hat keine Pläne, zu einer dezentralen Architektur zu wechseln»: Meredith Whittaker.

In Grossbritannien steht ein ähnliches Gesetz namens «Online Safety Bill» kurz vor der Verabschiedung. Wie wird Signal darauf reagieren?
Wir sind zu diesem Thema in den britischen Medien gerade sehr präsent. Und wir werden alles tun, was wir können, um in Grossbritannien im App-Store verfügbar zu bleiben. Wir haben auch Proxy-Server im Iran eingerichtet, als die Regierung Signal dort sperren liess. Doch wenn wir wegen dieses britischen Gesetzes gezwungen werden, unsere Datenschutzversprechen zu brechen und damit unsere Daseinsberechtigung zu verlieren, müssen wir uns aus dem britischen Markt zurückziehen.

Reden wir nun über den «Digital Markets Act» der EU, der ebenfalls Messenger-Apps betrifft. Zurzeit sind Signal, Whatsapp und Telegram alles «geschlossene Gärten», die nicht untereinander kommunizieren. Neu könnten aber mit dem «Digital Markets Act» kleinere Messenger-Apps wie Signal die «Grossen» wie Whatsapp von Meta oder iMessage von Apple dazu zwingen, interoperabel zu werden.
Genau.

Damit sollte dieselbe Kommunikation wie bei E-Mail möglich werden: Jemand mit einer Gmail-Adresse kann jemandem mit einer Proton-Adresse direkt eine Mail schicken. Welche Position hat Signal dazu?
Ich verstehe diesen grossen politischen Wunsch nach Standardisierung

wie bei der E-Mail. Ich befürchte aber, dass das nicht funktionieren wird. Wir nutzen bei Signal ein starkes technisches Protokoll, das den Inhalt der Nachrichten schützt, die Ende-zu-Ende-Verschlüsselung. Wir sind nie im Besitz dieser Schlüssel, die Geräte unsere Nutzerinnen tauschen diese in der Kommunikation miteinander aus.

Whatsapp wendet dieses Protokoll auch an.
Ja. Aber wir verwenden darüber hinaus weitere kryptografische Techniken, die Metadaten und Daten von Gruppenchats schützen. Wir leisten also in diesem Bereich mehr als das, was die Grossen wie Meta im Bereich des Datenschutzes anbieten. Und so stellt sich uns die Frage: Wie würden wir überprüfen, ob Whatsapp tatsächlich die Datenschutzstandards erhöhen wird, wenn wir mit ihnen zusammenarbeiten?

Besteht die Gefahr, dass der Datenschutzstandard von Signal quasi «nach unten» nivelliert wird?
Ja. Die Frage ist: Dürften wir die technischen Systeme von Whatsapp und Co. im Hintergrund anschauen, um sicherzustellen, dass bei einem messengerübergreifenden Nachrichtenversand nicht doch irgendwelche Metadaten gespeichert werden? Und werden Gatekeeper wie Apple und Meta ihre Datenschutzstandards an diejenigen von Signal anpassen? Das wäre unsere ultimative Bedingung für Interoperabilität. Und ja, ich fürchte, wir werden deswegen auch noch in zehn Jahren keine Kommunikation zwischen Messengersystemen haben.

Könnte man sagen: Es ist ein gut gemeintes EU-Gesetz für mehr Wettbewerb, das aber ohne Abstriche beim Datenschutz kaum umsetzbar ist?
Es liegt nicht daran, dass die EU-Kommission keine guten Absichten hat. Gute Absichten allein erfüllen aber noch keine Anforderungen für technische Spezifikationen der Internet Engineering Task Force (IETF). Dafür müssen technische Standards definiert werden. Und ich habe bisher noch nichts gesehen, was für unsere Ansprüche geeignet wäre.

Sie sind also skeptisch, was die Umsetzung des «Digital Markets Act» angeht?
Ja. Und ausserdem: Angenommen, es läuft gut, der Facebook-Messenger, Whatsapp, Signal und iMessage kommunizieren alle miteinander. Die Sorge von Signal ist, dass dies die Marktmacht der Grossen wie Whatsapp und iMessage noch mehr zementieren wird. Die Nutzerinnen würden ja dort verbleiben. Denn warum sollten sie zu uns wechseln, wenn wir ihnen nicht mal den höheren Datenschutz von früher anbieten würden? Wir würden uns selbst kannibalisieren.

Zum Schluss: Ich habe Nutzerinnen des dezentralen Netzwerks Mastodon gefragt, was sie von der Präsidentin von Signal wissen möchten. Jemand fragte: Sieht Meredith Whittaker die Zukunft der Internetwelt bei zentralisierten Netzwerken wie Twitter, Instagram oder eben auch Signal? Oder bei dezentralen Netzwerken wie Mastodon, mit unterschiedlichen Servern und Instanzen, die aber untereinander kommunizieren?
Diese Frage musste natürlich von Mastodon kommen. *(lacht)* Ich liebe das Tech-Ökosystem von Mastodon. Aber was uns angeht, ist klar: Signal hat keine Pläne, zu einer dezentralen Architektur zu wechseln. Ich weiss, dass die Mastodon-Nutzerschaft sich eine andere Antwort von mir wünscht. Doch für die Zentralisierung unseres Dienstes gibt es eine Reihe von Gründen.

Nämlich?
Wir haben viele Nutzerinnen bei Signal, die auf uns angewiesen sind. Das

heisst: Wir müssen eine robuste Infrastruktur haben und immer verfügbar sein. Wir können keine Echtzeit-Videocalls auf Signal ohne Content Delivery Networks durchführen. Signal wird in risikoreichen Ländern von Aktivistinnen und Oppositionellen eingesetzt, die auf unsere App angewiesen sind. Das dürfen wir nie vergessen. Es muss immer funktionieren, 24/7. Das ist der Hauptgrund.

Wäre denn ein stabiler Betrieb von Signal mit einer dezentralen Architektur in jenen Regionen so viel schwieriger aufrechtzuerhalten?
Ja, solche verteilten Architekturen sind schwer zu aktualisieren. Wenn du mit so einem System arbeiten würdest, müsstest du jede Person kennen, die einen Server für Signal betreibt, und dieser Person sofort einen Patch schicken. Dann merkst du vielleicht mal: Oh, die Hälfte unserer Community weltweit hat diesen Patch gar nicht bekommen. Und das wirkt sich dann wieder auf die IT-Sicherheit aus.

Ein dezentrales System von Signal ist also eine Utopie?
Nochmals: Ich verstehe den Wunsch, ich selbst mag die zentralisierte Internetwelt und die Big-Tech-Unternehmen mit ihren riesigen Datenbanken auch nicht. Ich kämpfe auch dagegen an, aber unser Ziel bei Signal ist es, den Menschen eine App zur Verfügung zu stellen. Und nicht, «am korrektesten» zu sein. Zudem: Dezentralisierte Infrastruktur bedeutet nicht automatisch auch dezentralisierte Macht. Ich glaube, das wird in der Kryptowelt sehr deutlich, in der oft gesagt wird: «Oh, Bitcoin ist eigentlich ein dezentralisiertes System, bei dem alle mitmachen können und alle profitieren.» Dann sage ich: «Ja, aber schau mal, wie viel Geld und wie viel Macht die reichen Winklevoss-Zwillinge mit ihren Kryptobörsen aufgebaut haben.»

Also braucht es Ressourcen, um überhaupt ein Player dieses dezentralen Systems sein zu können.
Jeder kann einen Server aufstellen, aber man muss schon eine Menge mitbringen, um so was stabil und zuverlässig für viele Menschen betreiben zu können. Und wer viel Geld hat, kann viele Serverinfrastrukturen anbieten und ein mächtiger Player in einem dezentralen System werden. Dieses Angebot würde dann wieder massenhaft Nutzerinnen anlocken. Das führt dann aber genauso zu einer Machtakkumulation von wenigen.