

«Das dringlichste Problem ist die Transparenz»

Die Möglichkeiten von Chatbot-Systemen werfen ethische Fragen auf, über die sich die Gesellschaft erst klar werden muss. Die EP/PL sprach mit dem Ethiker Jean-Daniel Strub.

BETTINA GUGGER



Engadiner Post: Herr Strub, ist die Gesellschaft auf Chatbot-SCHWERPUNKT Systeme vorbereitet?

Jean-Daniel Strub: Ich habe den Eindruck, die Gesellschaft ist besser darauf vorbereitet, als es bei früheren Technologiesprüngen der Fall war. Im Rahmen meines Hochschulunterrichts habe ich mit Studierenden ChatGPT angeschaut, auch einmal mit einer Gruppe von zumeist über 80-Jährigen. Ich nehme eine grosse Faszination, Interesse und Neugierde wahr, aber auch eine kritische Distanz. Das werde ich positiv. Nicht vorbereitet war man aber auf die Geschwindigkeit der technologischen Entwicklung. Gewisse Branchen haben bereits begonnen, selbstverständlich mit dem Tool zu arbeiten. Gewisse Fertigkeiten, die wir uns aneignen müssen, sind vergleichbar mit jenen, die wir uns bei der Nutzung von Google zu eigen machen mussten, das zeigt sich auch daran, dass jetzt bereits sogenannte «Prompt Engineers» eingestellt werden, also Personen, die besonders gut darin sind, die richtigen Abfragen für herausragende Resultate zu tätigen. Aber im Gegensatz zu Google gibt ChatGPT die Quellen der Informationen nicht an, was eines der Probleme ist.

Italien hat GPT4 bereits verboten, wie schätzen Sie dieses Vorgehen ein?

Das ist ein interessanter Schritt, der Nachahmung finden könnte. Italien nimmt auf den Datenschutz Bezug, da es zur Grundlage des Tools gehört, mit Daten der Nutzenden zu arbeiten, was deren Einwilligung bedarf. Es gilt immer abzuwägen, ob uns es wert ist, eine Dienstleistung zu beziehen und im Gegenzug Daten zur Verfügung zu stellen. Nicht dienlich scheint mir aber, wenn einzelne Nationen unkoordiniert vorgehen.

Zwingt uns die Gesellschaft nicht irgendwann dazu, solche Dienste auch zu benutzen?

Dazu habe ich eine klare Haltung, egal, welche Dienste es betrifft: Menschen, die ein digitales Tool nicht nutzen wollen, müssen in der Gesellschaft auch Platz haben. Es ist jedoch nicht die Aufgabe der Hersteller, für diese Frage eine Lösung zu liefern, sondern hier ist die Gesellschaft selbst in der Pflicht.

Sollte sich eine internationale Ethikkommission um Fragen rund um neue Technologien wie ChatGPT kümmern?

Der offene Brief, den KI-Expertinnen und -Experten Ende März im Hinblick auf die Lancierung von ChatGPT4 pu-

Jean-Daniel Strub

... gründete 2018 zusammen mit Johan Rochel das Beratungsunternehmens ethix, das Unternehmen und Stellen der öffentlichen Hand zu ethischen Themenkomplexen rund um Digitalisierung und Innovation berät.

Strub ist promovierter Theologe. Er war bis 2017 fünf Jahre im ETH-Rat tätig und von 2008 bis 2012 sowie während der Corona-Pandemie Geschäftsleiter der Nationalen Ethikkommission. (bg)



Jean-Daniel Strub verfügt über jahrelange Erfahrungen aus seiner Mitarbeit bei Ethikkommissionen.

Foto: z. Vfg

bliziert haben, will die öffentliche Diskussion voranbringen. Das ist zu begrüssen. Der Brief rückt allerdings in meinen Augen die dystopischen Vorhersagen zu sehr in den Vordergrund, sodass die aktuellen Risiken in den Hintergrund geraten könnten. Bevor wir ein Gremium wie eine Ethikkommission als Lösung propagieren, müssen wir uns über die Natur des Problems klar werden: Geht es darum, dass Unternehmen, die öffentliche Hand und politische Systeme ganz offensichtlich miteinander im Wettbewerb stehen oder um die Gefahr, von der Maschine überholt zu werden? Oder geht es um den Machtgewinn einzelner Unternehmen bezüglich der Beeinflussung von Meinungen und Verbreitung von Falschinformation, insbesondere auch von Bildmaterial? Eine Ethikkommission einzusetzen, die nicht weiss, auf welches Problem sie reagieren soll, ist wenig wirkungsvoll, es sei denn, sie hat die nötige Zeit, beim Sortieren der Fragestellungen behilflich zu sein.

Was ist aus Ihrer Sicht das dringlichste Problem?

Das dringlichste Problem ist die Transparenz, zu wissen, was mit welchen Daten und aus welchem Grund passiert. Andererseits geht es um die Bewusstwerdung, dass die Tools vorhandene Information zusammenfügen und man

die Informationen daher immer mit Blick auf den Wahrheitsgehalt der Aussagen überprüfen muss. Schliesslich gilt es auch zu fragen, ob man die Unterstützung im einzelnen Fall überhaupt braucht, um zu einer Erkenntnis zu gelangen.

Wie beurteilen Sie die negativen Auswirkungen von Chatbot-Systemen auf die Psyche des Einzelnen, insbesondere die Gefahr der Vereinsamung?

Ich sehe keine belastbaren Hinweise für eine Entwicklung, die dazu führt, dass Menschen durch die Nutzung solcher Systeme echte soziale Beziehungen vernachlässigen. Ähnliche Befürchtungen gibt es bei gewissen sozialen Medien oder beim Gaming. Auch dort ist das Problem vor allem, dass sich Menschen unter Druck gesetzt fühlen, nicht in erster Linie, dass sie soziale Kontakte aufgeben.

Wie schätzen Sie die Gefahr des Verlustes von Arbeitsplätzen ein?

In bestimmten Bereichen ist ein Form- und Strukturwandel denkbar; künftig muss nicht mehr jede Arbeit von Menschen gemacht werden. Gewisse Jobs könnten unter Druck geraten, namentlich in der Kreativ- und der Medienbranche. Aber gerade diese Branchen sind aufgrund der Digitalisierung ja schon länger in Umwälzung begriffen. Bereits 2013 gab eine Oxford-

Studie, die prophezeite, dass 50 Prozent der (US-amerikanischen) Arbeitsplätze infolge der Digitalisierung wegfallen würden. Das ist nicht geschehen, es wurden eher mehr Jobs geschaffen. Wichtig zu beachten ist aber, dass jetzt auch klassische Mittelstands- und Bürojobs unter Druck kommen. Das ist anders als bei früheren Technologieschüben, bei denen vor allem Handarbeit rationalisiert wurde.

Welche Massnahmen können Entwickler von Chatbot-Systemen ergreifen, um sicherzustellen, dass keine diskriminierenden oder rassistischen Algorithmen entstehen?

Das ist eine gute Frage, die nicht nur ChatGPT im Speziellen betrifft. In der Vergangenheit liessen verschiedene Algorithmen diskriminierende Tendenz erkennen und mussten korrigiert werden. Ein zentraler Aspekt ist auch hier die Transparenz: Entwickler müssen auf die Qualität der Trainingsdaten Acht geben, damit Diskriminierungen, die man in der realen Welt vorfindet, nicht reproduziert werden. OpenAI, die Firma, die ChatGPT entwickelt hat, legt durchaus ein Augenmerk darauf.

Im Gegenzug müssen wir davon ausgehen, dass Militärs und Geheimdienste die Technologie für ihre Zwecke missbrauchen ...

Umso mehr müssen wir uns auf ethische Standards einigen, welche die Unternehmen verpflichten. Als Gesellschaft haben wir den Anspruch, dass Rahmenbedingungen entstehen, die gewisse Formen von Missbrauch verhindern, nicht nur spezifisch für ChatGPT.

Wie schätzen Sie die Gefahren von Manipulation und Propaganda ein?

Die sind sehr real und sicher noch ausgeprägter mit den neuen Entwicklungen. Die Bildpresse ist da noch exponierter; die KI-generierten Fotos der angeblichen Verhaftung von Donald Trump oder der fiktive Kniefall von Putin sind Beispiele dafür, wie schnell solche Bilder im Informationsstrudel ein Eigenleben entwickeln. Die Hersteller müssen in die Pflicht genommen werden und künstlich erzeugte Realitäten als solche kennzeichnen.

Wie beurteilen Sie die Machtkonzentration der Tech-Giganten und global tätiger Unternehmer wie Bill Gates oder Elon Musk, deren Einfluss weit über die Tech-Branche hinausgeht?

Das ist in der Tat ein wichtiges Thema. Auch diese Frage stellt sich jedoch mit ChatGPT nicht neu, denn das Internet hat schon seit Beginn eine «Winner takes it all»-Dynamik bedient. Wenn man allerdings sieht, wie viel Kapital seit der Veröffentlichung von ChatGPT im November in die Entwicklung immer leistungsstärkerer KI-Tools fliesst, wird noch einmal deutlich, wie gross die wirtschaftlichen Interessen auf diesem Gebiet inzwischen sind und wie wichtig es dementsprechend ist, die Akteure in diesen Branchen wirkungsvoll für ein verantwortungsvolles Handeln in die Pflicht zu nehmen. Dabei ist sicher auch eine rechtliche Regulierung unerlässlich. Eine Frage ist, wie die Geschwindigkeit der technischen Entwicklung mit der natürlichen Langsamkeit demokratischer Prozesse in Einklang gebracht werden kann.

Werden KI-Systeme irgendwann so weit sein, dass sie selbstständig ethische Entscheidungen fällen?

An diesem Punkt sind wir glücklicherweise noch nicht. Algorithmen können uns aktuell bei der Entscheidungsfindung unterstützen, sie können gewisse Wahrscheinlichkeiten präziser voraussagen als eine menschliche Kalkulation, aber natürlich dürfen sie uns die Entscheidung nicht diktieren. Bei den heute diskutierten Anwendungen sind wichtige Merkmale für die Entscheidungsfähigkeit nicht gegeben. Sie haben zum Beispiel kein eigentliches Verständnis einer Situation und verfügen auch nicht über das, was wir Intuition nennen.

Inwiefern können Chatbot- und Robotik-Systeme zur Rechenschaft gezogen werden?

Die Systeme sind nicht autonom handelnde. Sie haben keinen eigenen Willen, auch keine Form von eigener Intelligenz, in dem Sinne, dass sie etwas Neues schaffen, das über das hinausgeht, was sie – in unvorstellbar grossen Datenmengen – vorfinden und auswerten. Daher können die Systeme selbst keine Verantwortung übernehmen. Wer diese Verantwortung trägt, muss mit Blick auf die Anwendungen aber zwingend geklärt werden. Verantwortlich kann einerseits der Anwender sein, wenn er das System nicht ordnungsgemäss einsetzt. Vor allem aber tragen die Hersteller Verantwortung, und zwar sowohl für die Sicherheit der Systeme als auch für die gesellschaftlichen Auswirkungen, die mit ihnen einhergehen. Wir müssen darauf hinwirken und hoffen, dass sie diese Verantwortung bei ChatGPT und Co. noch viel stärker wahrnehmen.