Sarah Spiekermann: "Wir vertrauen der KI zu sehr"

Anwendungen wie ChatGPT klingen klug, sind aber unzuverlässig, sagt Sarah Spiekermann. Die WU-Professorin für Wirtschaftsinformatik über Regulierung, Zulassungsprozesse und Misstrauen.

Thema: Künstliche Intelligenz (KI)  von **Martina Bachler** / Interview aus trend. Edition vom 27.4.2023veröffentlicht am **23.5.2023**

trend: Künstliche-Intelligenz-Anwendungen wie ChatGPT oder Midjourney wachsen rasant. Ist es schon zu spät, diese Programme zu regulieren?
Sarah Spiekermann: Nein, denn die technische Entwicklung wird noch Jahrzehnte weitergehen. Die heute gesetzten Maßnahme werden das gegenwärtige und das zukünftige Design der Technologie bestimmen. KI ist außerdem schon zum Teil reguliert. Würden die Gesetze befolgt, hätten sie sogar Wirkung.
Welche Gesetze meinen Sie?
In der Datenschutzverordnung (DSVGO) der EU ist etwa "Privacy by Design" vorgeschrieben. Software müsste schon jetzt durch ihr Design dafür Sorge tragen, dass sie den Datenschutz nicht verletzt. Auch das Urheberrecht gibt es ja bereits. Nun aber wird diskutiert, inwieweit KI-Programme es verletzen. Und natürlich ist es nicht zu spät für weitere Regulierung.
Warum ist sie zum Beispiel nötig?
Zur Gefahr wird etwa ChatGPT, weil seine Antworten ein Sammelsurium aus Information und Fantasie sind. Diese KI reiht einfach Wort an Wort, sie klingt intelligent, ist es aber nicht. Sie liegt oft richtig, nur weiß man nicht, wann. Sie kann etwa sehr verzerrende Antworten zu Personen geben, was aber deren Rechte und Würde verletzt.

Wir neigen dazu, künstlichen Intelligenzen oder Maschinen mehr zu glauben als Menschen.

Einem australischen Politiker dichtete ChatGPT etwa sexuelle Belästigung an.
Gefährlich ist das deshalb, weil wir dazu neigen, künstlichen Intelligenzen oder Maschinen mehr zu glauben als Menschen. Das ist ein echtes Problem. Im Silicon Valley gibt es diese süffisante Bezeichnung einer "Highly Paid Personal Opinion", also "Hochbezahlte persönliche Meinung." Das ist ein ziemlich verachtender Ausspruch über das Urteil von Menschen mit Erfahrung. Aber so ist es: Wir glauben heute eher der Maschine.
Braucht es für KI also eine Zulassung wie für Medikamente?
Im Zuge der KI-Verordnung wird in Brüssel auch das gerade diskutiert. Der bisherige Ansatz sieht so aus, dass Systemtypen, bei denen man ein gewisses Risiko für die Bevölkerung wittert, eine Zulassung durchlaufen müssen. Das soll nun auch Programme wie ChatGPT inkludieren. Allerdings wurde kürzlich in einer Präsentation der EU-Kommission gesagt, dass der gegenwärtige Verordnungsentwurf nur 15 Prozent der KI-Systeme einer Risikoprüfung unterzieht. Aus meiner Sicht ist dieses Risikomodell nicht optimal. Erst eine Zulassungsprüfung selbst kann wirklich determinieren, welche KI als Gefahr einzustufen ist und welche nicht. China geht da nun viel strenger vor.
China hat gerade einen Gesetzesentwurf vorgelegt, der etwa ein Sicherheits-Assessment und auch Transparenz der Algorithmen voraussetzt.
Er sieht vor, dass Unternehmen, die die KI herstellen, auch für die Richtigkeit der berechneten Inhalte geradestehen. Die USA und Brüssel hingegen scheinen sich schwerzutun, KI-Provider dafür haften zu lassen. Man hat Angst, technologisch durch Regulierung im Wettbewerb zurückzufallen. Dabei wird die inhaltliche Verantwortung für Plattformen etwa von klassischen Medien lange gefordert, aber die Lobbyarbeit der IT-Konzerne ist zu erfolgreich.

Unternehmen müssen sich sehr genau überlegen, ob sie eine so fehleranfällige Technologie wie diese neuen KI-Sprachmodelle einsetzen.

KI-Anbieter sollen also etwa für Falschinformationen haften?
Wenn Journalisten heute auf der Basis ihrer Recherchen falsche Angaben machen, dann steht der Herausgeber dafür grade. Analog sollte, wenn eine KI nun in Zukunft eine Reise zusammenstellt, und dann gibt es den behaupteten Flug nicht, auch der Hersteller der KI dafür haften bzw. derjenige, der die Dienstleistung dem Endkunden zur Verfügung stellt.
Unternehmen sollen also vorsichtig sein?
Egal, wie die Regulierung aus Brüssel aussehen wird, Unternehmen müssen sich sehr genau überlegen, ob sie eine so fehleranfällige Technologie wie diese neuen KI-Sprachmodelle einsetzen. Fehler haben Opportunitäts-und Folgekosten. Die Gefahr besteht jedoch, dass Unternehmen eine 80-prozentige Richtigkeit in vielen Anwendungsbereichen genügt. Die Ungenauigkeit bei den neuen KI-Sprachmodellen finde ich sehr bedauerlich. Mir würden nämlich viele Services einfallen, bei denen ein ChatGPT helfen könnte -wäre es denn zu 99 Prozent zuverlässig. Mein Rat ist daher, dass man besser wartet, bis die Technik reif ist.
Sie haben einen Standard für ethische KI-Systeme mitentwickelt. Was schreibt er vor?
Das Ziel des ISO-Standards ISO/IEEE 24748-7000 ist es, sozial verträgliche, humane KI-Systeme zu entwickeln. Diese müssen mehrere Prozesse durchlaufen, um sicherzustellen, dass man sie verantwortungsvoll auf den Markt bringen kann. Bei so einem Verfahren könnte herauskommen, dass die aktuell besprochenen KI-Modelle aber zum Beispiel die Anforderungen an Wahrheitsgehalt technisch gar nicht erfüllen können, selbst wenn man in dem Zulassungsverfahren feststellt, wo etwa ChatGPT nachbessern müsste. Möglicherweise gibt es aber Begleitmaßnahmen, die es dann vertretbar machen, das System auf den Markt zu bringen.

Wäre die KI verlässlich, könnte sie völlig ändern, wie wir das Internet nutzen.

Wie läuft dieses Verfahren ab?
Im ersten Schritt versucht man, auf Basis der Datenflüsse und Softwaremodule zu verstehen, welche Daten die Maschine woher bezieht und wie sie diese verarbeitet. Dann muss man verstehen, in welchem Kontext das System wie genutzt wird. Je nachdem muss man es unterschiedlich beurteilen in Bezug auf die Möglichkeit, Rechte und Werte Einzelner zu verletzen. Danach muss man sich anschauen, welche Wertkonsequenzen die Technologie hat. Dabei geht es gar nicht so sehr um Datenschutz, sondern um Werte wie Würde, um Wahrheit, aber auch Autorenschaft. Im dritten Schritt formuliert man die technischen und organisatorischen Anforderungen an das System so, dass wichtige Werte nicht verletzt und positive Werte sogar gefördert werden können. IEEE hat eine Zusammenarbeit mit TÜV Süd vereinbart und bildet jetzt Coaches aus, die Unternehmen bei diesen Projekten begleiten bzw. sogar zertifizieren können. Ich selbst habe den ISO-Standard anhand von Fallstudien mit der Unicef, mit BMW und mit Start-ups ausprobiert.
Erleben wir gerade einen entscheidenden Moment der Technikgeschichte?
Wäre die KI verlässlich, könnte sie völlig ändern, wie wir das Internet nutzen. Es gäbe dann einen wirklich disruptiven Moment mit vielen Folgen. Die Gefahr ist aber groß, dass wir der jetzt vorgestellten Art von KI-Service gutgläubig vertrauen, weil wir als Menschheit lange auf den Moment gewartet haben, in dem Maschinen so intelligent sind, wie ChatGPT klingt. Deshalb ist es so wichtig, dass Firmen verstehen, womit sie es derzeit zu tun haben.
Wie optimistisch sind Sie, dass die Regulierung dieses Mal früh funktioniert?
Das Formulieren von Gesetzen ist heute stark von Unternehmen beeinflusst, ihre Durchsetzbarkeit befindet sich in einer Krise. Ich begrüße, dass die EU-Kommission die neue Verordnung anstößt. Entscheidend wird aber sein, wie weise die breite Masse an Unternehmen mit dieser Innovation umgeht.

ZUR PERSON
Sarah Spiekermann, 49, leitet seit 2009 das Institut für Informationssysteme &Gesellschaft an der Wirtschaftsuniversität Wien.