Technologieentwicklung: Wie kommt Ethik in die KI?



Künstliche Intelligenz (KI) verspricht viele nutzbringende Anwendungen. Inwieweit diese realisiert werden, hängt maßgeblich davon ab, ob die Menschen Vertrauen in die Technologie setzen. Die Plattform Lernende Systeme hat einen Leitfaden zur verantwortungsbewussten Entwicklung und Anwendung von KI-Systemen erstellt. Jessica Heesen, Leiterin des Forschungsschwerpunkts Medienethik und Informationstechnik an der Universität Tübingen sowie Leiterin der Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik der Plattform Lernende Systeme, erläutert im Interview maßgebliche Kriterien.

Im europäischen Diskurs über Künstliche Intelligenz (KI) spielen ehische Werte und Prinzipien eine wichtige Rolle. Eine große Mehrheit der Deutschen wünscht sich einer aktuellen Umfrage des Bitkom zufolge sichere KI-Systeme und erwartet, dass diese besonders gründlich geprüft werden. Was aber beinhaltet eine verantwortungsbewusste Entwicklung von KI-Systemen? Antworten darauf liefert ein Leitfaden der Plattform Lernende Systeme, einem vom Bundesministerium für Bildung und Forschung (BMBF) initiierten Expertennetzwerk. Er greift den in der KI-Strategie der Bundesregierung formulierten Ansatz "Ethics by, in and for Design" auf und will Entwicklerinnen und Entwicklern, Anbietern, Nutzern und Betroffenen von KI-Systemen eine Orientierung geben. Im Interview erläutert Mitautorin Jessica Heesen, Leiterin des Forschungsschwerpunkts Medienethik und Informationstechnik am Internationalen Zentrum für Ethik in den Wissenschaften (IZEW) der Universität Tübingen sowie Leiterin der Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik der Plattform Lernende Systeme, wichtige Kriterien einer verantwortungsbewussten KI-Entwicklung.

Frau Heesen, welchen Prinzipien sollte eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen folgen?

Jessica Heesen: Auf prinzipieller Ebene soll die KI-Entwicklung der Gesellschaft dienen und nicht dazu führen, dass neue technische oder wirtschaftliche Zwänge entstehen, die ethische Normen des Zusammenlebens verletzen oder positive Entwicklungen beschränken. Das ist damit gemeint, wenn in vielen politischen Dokumenten und Reden von "KI für den Menschen" die Rede ist. Und das hat auch schon der Philosoph Theodor W. Adorno 1953 zum Ausdruck gebracht: "Es gibt keine technologische Aufgabe, die nicht in die Gesellschaft fällt". Technikentwicklung beschränkt sich somit nicht auf die technische Lösung an sich, sondern ist elementarer Baustein einer menschenwürdigen Gesellschaft. Dementsprechend muss sich die Entwicklung von KI als eine der modernen Schlüsseltechnologien an Leitwerten wie Diskriminierungsfreiheit, Schutz der Privatheit, Zuverlässigkeit oder Transparenz orientieren.

Generell ist es wichtig, die Handlungsroutinen und Handlungsempfehlungen, die durch KI-Systeme vorgegeben werden, nicht als alternativlos ("Sachzwang") anzuerkennen. Wie in andere technische Produkte auch, sind in KI-Software bestimmte, veränderbare Zwecke und Präferenzen "eingeschrieben", die bestimmten Gruppen und Individuen nutzen, anderen aber schaden können.

Was bedeutet das konkret für die Entwicklung und Anwendung von KI?

Jessica Heesen: Wir müssen hier zwei Ebenen unterscheiden: Zum einen die Handlungsmöglichkeiten von Einzelnen oder Gruppen bei der Nutzung von KI und zum anderen die politischen und rechtlichen Rahmenbedingungen, die für einen wertorientierten und sicheren Einsatz von KI gesetzt werden.

So lassen sich etwa algorithmische Entscheidungssysteme mit KI-Komponenten dazu nutzen, um über Sozialleistungen zu entscheiden. Die Vereinten Nationen sprechen hier - durchaus kritisch - von einem "Digital Welfare State". KI-Systeme können zum Beispiel - wie in Indien bereits praktiziert - über die Vergabe von Lebensmittelgutscheinen entscheiden. Oder aber potenziellen Sozialbetrug aufspüren. Die Nutzung eines Programms namens SyRI, mit dem das niederländische Sozialministerium Menschen identifizierte, die möglicherweise zu Unrecht Arbeitslosen- oder Wohngeld bezogen hatten oder dies künftig tun könnten, wurde mittlerweile gerichtlich untersagt.

Klar ist: Bei der Verwendung derart sensibler KI-Entscheidungssysteme ist ein hohes Verantwortungsbewusstsein gefragt. Die Verwaltungsangestellten vor Ort müssen die Funktionsweisen zumindest in groben Zügen verstehen, interpretieren und gerecht anwenden. Dazu gehört, sich in der Rolle des "menschlichen Letztentscheiders" auch gegen die Entscheidung eines KI-Systems stellen zu können. Die Behörde wiederum muss beim Einkauf eines solchen Systems auf Zuverlässigkeit und Diskriminierungsfreiheit achten. Die Entwicklerinnen und Entwickler stehen u. a. in der Verantwortung, die Trainingsdatenqualität sicherzustellen und die Systeme geregelt nach diskriminierenden Faktoren zu untersuchen.

Welche Implikationen ergeben sich daraus für eine mögliche Regulierung von KI-Systemen?

Jessica Heesen: Der Regulierungsbegriff umfasst ein ganzes Spektrum von Möglichkeiten, um KI-Systemen eine Form zu geben, die eine Gesellschaft technisch voranbringt, zugleich aber im Einklang steht mit den Wertvorstellungen liberaler Rechtsstaaten. Deshalb betont die EU-Kommission immer wieder den Stellenwert einer wertorientierten Entwicklung von KI. Generell gibt es dafür verschiedene Ansätze: Schwache Regulierungsformen sind beispielsweise Ethikkodizes von Fachgesellschaften oder Unternehmen. Starke Regulierungsformen sind rechtliche und staatliche Vorgaben in Form von Zertifizierungen von KI. In all diesen unterschiedlichen Regulierungsformen sind ethische Wertvorstellungen grundlegend und müssen in einem weiteren Schritt für die die jeweiligen Anwendungskontexte ausformuliert werden.

Die Plattform Lernende Systeme hat einen Ethik-Leitfaden entwickelt, der für Entwicklerinnen und Entwickler, Anbieter, Nutzer und Betroffene einen ethisch reflektierten Entwicklungs- und Anwendungsprozess von KI-Systemen skizziert. Er benennt Prinzipien wie Selbstbestimmung, Gerechtigkeit und Schutz der Privatheit, aus denen die Kriterien für die Regulierung von KI-Systemen großteils hervorgehen. Wir möchten damit eine Beitrag zur öffentlichen Diskussion leisten und eine Impuls für weiter Debatten geben.

Das Whitepaper ?Ethik-Briefings ? Leitfaden für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen? der Plattform Lernende Systeme steht hier zum kostenlosen Download bereit:
https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG3_Whitepaper_EB_200831.pdf