



Jetzt günstig versichern!

Kritik am Ethikrat

KI kann menschliche Verantwortung nicht ersetzen? – Doch!

Eine Kolumne von Sascha Lobo

Der Deutsche Ethikrat romantisiert in seiner Stellungnahme zur künstlichen Intelligenz den Menschen. Auf zwei konkrete Gefahren der Technik geht er aber kaum ein. Eine davon demonstriert unser Kolumnist.

22.03.2023, 16.29 Uhr

Artikel zum Hören • 8 Min



Alena Buyx präsentiert die Stellungnahme des Ethikrats: Allzu konservativer Blick auf künstliche Intelligenz Foto: ANNEGRET HILSE / REUTERS

Der Deutsche Ethikrat hat eine sogenannte **»Stellungnahme«** [vorgestellt](#), eine fast 300 Seiten lange Einlassung zum derzeitigen und auch zukünftigen Dauergroßthema künstliche Intelligenz. Vielleicht muss ich an dieser Stelle, vor Beginn meiner Kritik an der Stellungnahme, schon aus Transparenzgründen eine persönliche Verbundenheit vorwegschicken. Ich halte den Ethikrat für eine wichtige und gute Institution, nicht nur als Idee, sondern auch ganz konkret. 2015 durfte ich dort einen Vortrag über Big Data halten, an den sich eine Art Gasthörerschaft mit Diskussionsbeteiligung anschloss. Die intellektuelle Flughöhe und allgemeine Qualität der dortigen Diskussionen habe ich davor und danach kaum je wieder erlebt.



Sascha Lobo >

Foto: Urban Zintel

Jahrgang 1975, ist Autor und Strategieberater mit den Schwerpunkten Internet und digitale Technologien. 2019 erschien bei Kiepenheuer & Witsch sein Buch »Realitätsschock: Zehn Lehren aus der Gegenwart«. Gemeinsam mit Jule Lobo beschäftigt er sich im Podcast [»Feel the News – Was Deutschland bewegt«](#) mit aktuellen Debattenthemen.

Es ist kein Geheimnis, dass ich [künstliche Intelligenz](#) für [das wichtigste und wirkmächtigste digitalgesellschaftliche Thema unserer Zeit](#) halte. Und genau deshalb bin ich zwar strukturell froh um die Stellungnahme, weil sie eine Reihe von extrem relevanten Themen adressiert, wie etwa eine tiefe Betrachtung der künstlichen Intelligenz in der Medizin. Ich bin aber auch etwas enttäuscht – wegen des allzu konservativen Blicks auf künstliche Intelligenz. So konservativ, dass ich die Perspektive in Teilen für unrichtig halte.

ANZEIGE

Bei der Vorstellung der Stellungnahme sagte der ehemalige Kulturstaatsminister [Julian Nida-Rümelin](#) einen Satz, der die Essenz dieser konservativ-unrichtigen Annahme darstellt:
»KI-Anwendungen können menschliche Intelligenz, Verantwortung und Bewertung nicht ersetzen.« Meine kurze Antwort darauf wäre: Doch. Es geht sogar genau darum.

Absurdes Hohepriestertum um menschliche Fähigkeiten

Wir können gesellschaftlich gern über den Begriff der Verantwortung in diesem Kontext diskutieren. Aber wenn man davon ausgeht, dass Maschinen schon bald eine Reihe von Echtzeit-Entscheidungen treffen, die massive Folgen haben – ein selbstfahrendes Auto wechselt die Spur –, dann müssen Maschinen zwingend eine Verantwortung transportieren. Dass man am Ende darauf besteht, dass immer ein Mensch dafür geradestehen soll, ist davon zunächst unabhängig. Noch eindeutiger ist, dass die KI-Bewertung die menschliche ersetzen wird, zum Teil schon ersetzt und auch ersetzen soll. Das absurde Hohepriestertum um die ach so goldene menschliche Bewertungsfähigkeit im Alltag habe ich nie verstanden. Werte – ja, die soll und muss der Mensch festlegen und immer wieder ihre Einhaltung überprüfen, aber eine Bewertung auf Basis dieser Werte? Nein, das kann gut und gern Maschinensache sein.

ANZEIGE

Es gibt ein interessantes Beispiel zur Bewertung von Menschen. In der Stellungnahme wird natürlich die Problematik von Diskriminierung und Bias intensiv

beleuchtet, dazu auch die konkreten Schwierigkeiten, etwa im beruflichen Kontext oder im Sozialwesen. In der Diskussion in freier Wildbahn geht es meistens um Entscheidungen im Personalwesen und die Befürchtung, dass KI Frauen oder Minderheiten diskriminiere. Das ist auch eine reale Gefahr, aber der kann man begegnen. »Responsible AI« ist das Schlagwort, unter dem diese Anpassung des Instruments KI [gesellschaftlich und wissenschaftlich diskutiert wird](#), was in der Stellungnahme kurz angerissen wird. Dabei werden zum Beispiel die Entscheidungsgrundlage, die dazugehörigen Daten und die Entscheidungswege für oder gegen eine Person nachvollziehbar gemacht. Dadurch lässt sich diskriminierungsarme oder sogar -freie KI zumindest anstreben. Im Gegensatz dazu kann man aber sehr viel schwerer dem Kern der Diskriminierung durch menschliche Entscheider sinnvoll begegnen, denn die legen kaum je ihre wahren Motivationen offen. Es geht mir nicht darum, dass Menschen nichts mehr entscheiden sollen – sondern um eine offenere und ehrlichere Diskussion der aberwitzig vielen Unzulänglichkeiten der Menschen. Die sie selten wahrhaben wollen.



Das ist der Kern meiner Kritik an der Stellungnahme: Das Menschenbild, das der Diskussion zugrunde liegt, romantisiert an vielen Stellen den Menschen und seine menschliche Entscheidungsgewalt, etwa mit einem Satz wie: »Wollte man versuchen, den Menschen als Handlungssouverän zu ersetzen, käme es zu einer Diffusion oder vollständigen Eliminierung

von Verantwortung.« Ethik könnte prinzipiell auch »Mahnwesen« heißen, das liegt in der Natur dieser Disziplin. Aber diese große Mahnung verkennt aus meiner Sicht, dass der Mensch als Handlungssouverän oft genug komplett in den Quark greift und anschließend nicht einmal genau sagen kann, warum.

ANZEIGE



Sascha Lobo
Realitätsschock: Zehn Lehren aus der Gegenwart

Verlag: Kiepenheuer&Witsch

Seitenzahl: 400 Seiten

Für 21,70 € kaufen 

Bei Amazon bestellen 

Bei Thalia bestellen 

Bei Yourbook bestellen 

Produktbesprechungen erfolgen rein redaktionell und unabhängig. Über die sogenannten Affiliate-Links oben erhalten wir beim Kauf in der Regel eine Provision vom Händler. [Mehr Informationen dazu hier](#)

»Eliminierung von Verantwortung« ist sogar ein klassisches Muster der menschengemachten Bürokratie – der man maschinell entgegenwirken kann, weil die Maschine transparent sein kann und dadurch messbar wird, wenn man will. KI muss als Teil der Automatisierung begriffen werden, die jetzt auch geistige, intellektuelle, schöpferische Arbeit erfasst hat und in der Folge sogar zu einem eigenen Handlungssouverän werden kann. Der soll Regeln folgen und überprüfbar, sanktionierbar und veränderbar sein, natürlich, aber wenn die Regeln in erster Linie darin münden, dass am Ende ein Mensch den Knopf drücken soll, damit im Zweifel jemand schuld ist, sehe ich darin keinen ethischen Mehrwert, sondern Verantwortungskitsch.

ANZEIGE



Wenn Open Source zum Problem wird

Wenn es um die konkreten technischen Ausprägungen von KI geht, fehlen mir in der Stellungnahme des Ethikrats außerdem detailliertere Einlassungen zu zwei Gefahrenkomplexen, die für mich zu den mächtigsten gehören.

Der Erste ist der Minimalisierungsprozess: Inzwischen läuft eine dem 2020 von OpenAI veröffentlichten Sprachmodell GPT-3 ähnliche KI [sogar schon auf einem handelsüblichen Laptop](#), und zwar in Verbindung mit dem Open-Source-Fortschritt. Open Source ist eine wunderbare, digital essenzielle Sache, aber wenn alle Menschen frei verfügbare KI-Anwendungen auf ihren Laptops laufen lassen können, dann fällt jede Kontrollierbarkeit weg.

Die Regeln, die OpenAI wie fast alle anderen KI-Anbieter sich selbst gegeben hat und die immer wieder feinjustiert werden, sind dann umgehbar. Und damit auch der im Ethikratspapier geforderte regulatorische Rahmen. Dazu muss man wissen: Im Frühjahr 2022 veröffentlichten Wissenschaftler*innen einen [Artikel in der Fachzeitschrift »Nature«](#), für den sie eine KI mögliche biochemische Waffen entwickeln ließen. In sechs Stunden brachte das Modell 40.000 potenziell tödliche Moleküle hervor, darunter solche, die leicht herzustellen sind, und solche, die dem giftigsten aller Kampfstoffe, VX, ausgesprochen ähnlich sind. Sogar VX selbst designte die KI noch einmal, ohne von dessen Existenz zu wissen. Wenn derart machtvolle Instrumente für Einzelpersonen herunterladbar im Netz stehen und auf normalen Laptops laufen, ergibt sich ein völlig neues Szenario der Herausforderungen durch künstliche Intelligenz in Zukunft.

Deepfakes werden vom Ethikrat weitgehend ignoriert

Der zweite Gefahrenkomplex, der mir in der Stellungnahme fehlt, wird quasi ab sofort wirksam. In meinem Podcast »Feel the News – Was Deutschland bewegt« habe ich gemeinsam mit Jule Lobo besprochen, wie künstliche Intelligenz funktioniert und was genau die [Chancen und Risiken dieser Technologie sind](#). Deepfakes, also Fälschungen von Videos oder Tonaufnahmen, die sich sehr, sehr echt anhören und deshalb entsprechende Wirkungen entfalten können, stellen eines dieser Risiken dar. Obwohl sogar ein eigenes Kapitel von der Wirkung auf öffentliche Kommunikation und

Meinungsbildung handelt, hat der Ethikrat das Deepfake-Problem ausgelassen.

Ein simples, aber eindrucksvolles Beispiel, warum eine solche Betrachtung sinnvoll gewesen wäre, habe ich für diese Kolumne selbst erstellt, mithilfe einer Plattform namens Elevenlabs, die Stimmen anhand einer kurzen Probe synthetisieren kann. Das funktioniert auf Englisch schon sehr eindrucksvoll, hier zum Beispiel ein Statement eines synthetischen [Olaf Scholz](#), das ich einfach als Text bei Elevenlabs eingegeben habe:

Externer Inhalt

[Zur Datenschutzerklärung](#)

Noch kommen auf diese Weise produzierte Filme etwas kantig daher, aber man muss bei künstlicher Intelligenz grundsätzlich immer das stetige Verbesserungspotenzial mitdenken. Das meiste, was die KI nicht kann, kann sie *noch* nicht. **S**

Diskutieren Sie mit

[Feedback](#)

ANZEIGE



inTouch

So hat sich die Fürstin in den Jahren verändert

ANZEIGE



Autoguru

Das sind die schnellsten Polizeiautos der...

ANZEIGE



Joyn

Trash TV Neuzugang: "Forsthaus..."

Aktuell in diesem Ressort

Elektroschrott... Wie Microsoft Nutzer zum Update auf Windows 11 nötig

In zweieinhalb Jahren ist Schluss mit Updates für Windows 10. Doch die Hardwareanforderungen für einen Wechsel zu Windows 11 sind enorm hoch. Laut Experten werden ältere Rechner damit zu »Elektroschrott«.



Für iPhones, iPads und Macs Warum Sie Apples Highspeed-Updat...

Ermittlungen wegen... 288 Festnahmen...



Onlinedating Ü50 »Frauen in meinem Alter...