

Künstliche Intelligenz

Roboter können keine Moral

Warum das Gerede von superintelligenten, allmächtigen Maschinen nur ein großes Ablenkungsmanöver ist.

Von **Richard David Precht**

17. Juni 2020, 16:49 Uhr / Editiert am 20. Juni 2020, 18:35 Uhr / DIE ZEIT Nr. 26/2020, 18. Juni 2020 / [53 Kommentare](#) /

EXKLUSIV FÜR ABONNENTEN

ARTIKEL HÖREN

Für viele ist es nur eine Frage der Zeit: Ob in zehn, zwanzig oder fünfzig Jahren – irgendwann in diesem Jahrhundert wird die Superintelligenz geboren, eine "starke" künstliche Intelligenz (KI), die die kognitive Leistungsfähigkeit des Menschen in allen Bereichen weit übersteigt. Das Schicksal des Menschen liege dann in ihrer Hand. Die spannende Frage, so heißt es dann, sei nur, ob diese Superintelligenz gut oder böse sei – und ob es den Menschen rechtzeitig gelingt, dies sicher festzulegen.

Die KI-Branche befeuert solche Fantasien: Sie verkauft Anfangserfolge als Durchbrüche und schwelgt in dunklen Allmachtsfantasien. Doch die Wahrscheinlichkeit, dass in absehbarer Zeit eine Superintelligenz dem menschlichen Gehirn in allen Bereichen überlegen sein wird, ist verschwindend gering. Der Begriff "künstliche Intelligenz" wirkt auf viele IT-Experten wie ein Marketingtrick. Dass Maschinen überschaubare Aufgaben blitzschnell auszuführen lernen, ist das eine – eine allumfassende Superintelligenz etwas anderes. Wer glaubt, dass tote Rechenmaschinen Geist, Bewusstsein oder Gefühle entwickeln, hat nicht mal entfernt verstanden, was Geist, Bewusstsein und Gefühle sind.

Nein, künstliche Intelligenz [<https://www.zeit.de/2020/09/kuenstliche-intelligenz-experten-forschung-universitaeten-bundesregierung>] wird Homo sapiens nicht in den Schatten stellen. Und sie wird auch keinen eigenen bösen Willen entwickeln, um die Menschheit zu knechten oder auszulöschen. In der Evolution führte der lange Weg vom Überlebenswillen vieler zur Intelligenz sehr weniger Lebewesen. Die umgekehrte Richtung, dass aus Intelligenz ein

**RICHARD DAVID
PRECHT, 55**

55, ist Honorarprofessor für Philosophie an der Leuphana Universität Lüneburg und an der Musikhochschule Hanns Eisler. Er schrieb mehrere Bestseller. Dieser Essay basiert auf seinem neuen Buch *Künstliche Intelligenz und der Sinn des Lebens*, das in dieser Woche im Goldmann Verlag erscheint.

Wille erwächst, ist unmöglich. Auch menschliche Intelligenz besitzt keinen Willen – sondern unser Wille geht unseren höheren kognitiven Leistungen immer voraus. Und außerdem: Wer sagt denn, dass eine Superintelligenz ausgerechnet über Machtgelüste verfügte? Menschliches Machtstreben ist unseren animalischen Leibern erwachsen und ist beileibe nicht unser wichtigster Antrieb. Ein kurzer Blick an einen Strand genügt. Menschen haben so viele andere Gelüste als Macht – Nichtstun, Sex, Albernheit und Alkohol, kurzum: den Willen zur Ohnmacht!

Eine Superintelligenz, setzte sie eigenständig ihre Ziele, hätte wohl niemals biologisch-menschliche Bedürfnisse: keine Ressourcensicherung, keinen

Expansionswillen, keinen Machttrieb. Wenn überhaupt, gliche sie eher Sokrates als Jeff Bezos. Die Ängste, die Hightech-Gurus vor bösen Robotern schüren, sind bisweilen Selbstbespiegelung oder Paranoia – meist sind sie aber ein gigantisches Ablenkungsmanöver. Denn KI, falsch eingesetzt, kann Menschen, Kulturen und Demokratien durchaus gefährlich werden – allerdings nicht durch ihren bösen Willen.

Das Gefahrenpotenzial künstlicher Intelligenz entlädt sich nicht auf einmal. Es hat nichts Bombastisches, sondern steckt in unfreiwilligen Diskriminierungen, wie unlängst bei einer Gesichtserkennungssoftware zu sehen, die Schwarze benachteiligt; es steckt in programmierten Unzulänglichkeiten und unbeabsichtigten Folgen, eigentlich überall da, wo Maschinen über Menschen richten. Am gefährlichsten wird es dort, wo Computer moralische Entscheidungen treffen sollen. Etwa bei voll automatisiert fahrenden Autos, bei Waffen, bei Bewertungen über Straffällige oder bei der Besetzung von Arbeitsstellen.



Dieser Artikel stammt aus der ZEIT Nr. 26/2020. Hier können Sie die gesamte Ausgabe lesen.
[<https://premium.zeit.de/abo/diezeit/2020/26>]

Die Idee einer "ethischen Programmierung" ist eine Illusion. Denn bereits ein kleiner Blick auf die gravierenden Unterschiede zwischen menschlichem und maschinellm Problemlösen zeigt, dass sie nicht realistisch ist. Das, was sich Menschen von "ethischer Programmierung" versprechen, ist höchst widersprüchlich. Denn der Computer oder Roboter soll einerseits ethisch "gut" sein und andererseits in konkreten Dilemma-Situationen präzise funktionieren. Er soll moralisch und zielgerichtet zugleich sein.

Doch je mehr man von menschlicher Ethik und Moral versteht, umso deutlicher wird, dass dieser Anspruch unerfüllbar ist. Lernen kann man das bereits bei David Hume. Der schottische Philosoph analysierte im 18. Jahrhundert, dass unser Gefühl unserem Verstand vorausgeht. Er wird heute von der Psychologie umfassend bestätigt. Weil wir bestimmte gefühlte Überzeugungen, Haltungen und Weltanschauungen haben, suchen wir uns die dazu passenden vernünftigen Argumente. Abtreibungsgegner sind zumeist nicht deshalb gegen Abtreibung, weil vernünftige Argumente sie überzeugen. Sondern sie suchen sich mehr oder weniger vernünftige Argumente, weil sie gegen Abtreibung sind.

Tatsächlich richten wir uns in der Moral zumeist nach unserer Intuition und entscheiden nach Gefühl. Die menschliche Moral ist ein Ensemble von irgendwie nützlichen, instinktiven Handlungen und Haltungen. Vernunft allein gebiert dagegen keine Moral. Denn ohne soziale Gefühle wie Liebe, Zuneigung, Respekt, Mitleid, Furcht, Unbehagen, Ablehnung, Ekel, Scham weiß unsere Vernunft nicht, was Gut und Böse ist. Sollte man also daran denken, moralisch wichtige Intuitionen bei der "ethischen Programmierung" zu berücksichtigen? Ein schräger Einfall! Denn damit würde man etwas normieren, das nicht normierbar ist.

Menschen handeln nicht wie Maschinen

Über Jahrtausende hinweg hat sich der Mensch als das "Andere der Natur" definiert. Angesichts von Maschinen, die bestimmte kognitive Leistungen von Menschen übertreffen, müssen wir nun lernen, den Menschen als das "Andere der künstlichen Intelligenz" zu sehen. Nirgendwo zeigt sich das so deutlich wie in der Moral. Moral ohne Subjektivität ist keine Moral und Subjektivität ohne Moral keine Subjektivität. Moralische Urteile bestehen nicht nur aus Ergebnissen oder gar "Lösungen", sondern der Weg, der Akt der Entscheidung, ist selbst von größter Bedeutung. Wenn die antiken Griechen ihre Tugendethik umzusetzen versuchten, wollten sie sich dabei im moralischen Handeln selbst kultivieren. Die Anstrengung gehört ebenso zur Moral wie die Handlung. In Immanuel Kants Pflichtenethik ist das nicht anders: Wir sollen unserem guten Willen folgen und lernen, das, was wir sollen, auch tatsächlich zu wollen. Die Anstrengung und der Akt des moralischen Handelns gehören also zur Moral entscheidend mit dazu.

Die präziseste Analyse dieses Aktes lieferte zu Beginn des 20. Jahrhunderts der amerikanische Philosoph, Psychologe und Soziologe George Herbert Mead. Menschen sortieren ihre Umwelt, indem sie sie verdinglichen. Personen, Dinge, Einsichten, ja sogar ganze Weltanschauungen werden so zu fest umrissenen und bewerteten Objekten. Wie das Design dieser Objekte aussieht, bestimmen die subjektiven Umstände, durch die sie in mein Leben treten. Nur so komme ich zu festen Überzeugungen über Hunde, Donald Trump, Schalke 04 oder den

Kommunismus. Der Wert eines Objekts ist niemals davon zu lösen, wie er zustande kommt. Entscheidend ist am Ende nicht die Sache, sondern der Akt, in dem das Objekt geformt und damit bewertet wird.

Das hat gewaltige Folgen für die Moral. Auch hier geht es vor allem um das Situative und um den Kontext. Wirkliche moralische Entscheidungen werden eigentlich selten getroffen. Denn meistens wissen wir in sozialen Situationen ziemlich genau, was wir tun werden. Wir überlegen nur dann, wenn wir mit etwas überfordert sind und ein Konflikt entsteht: Wie bewerte ich das, was ich will, im Hinblick auf das, von dem andere wollen, dass ich es tue? In einer solchen Dilemma-Situation hilft mir meine Ausstattung mit Fürsorgeverhalten nicht weiter. Und es geht auch nicht darum, die eigene Lust oder den eigenen Schaden mit der Lust oder dem Schaden anderer abzugleichen. Viel wichtiger in Konfliktsituationen ist, die eigene Identität zu wahren. Wenn ich nicht weiß, was ich tun soll, treibt mich dies in eine Persönlichkeitskrise; ein unschöner Zustand, der mich dazu anhält, die Konsistenz meines Selbst wiederherzustellen. Menschen handeln also nicht wie Maschinen. Statt neutral auf die Umwelt reagieren sie immer auch auf ihre eigenen sozialen Erfahrungen. Sie beschäftigen sich mit ihrem Selbstkonzept. Was den Menschen zum Menschen macht, liegt nicht in apriorischen Festlegungen oder Programmierungen. Es ist die besondere Weise, wie wir mit unserer Umwelt kommunizieren. Wir reagieren nicht einfach auf sie, sondern wir konstruieren sie uns als unsere Welt.

Die menschliche Moral ist irrational, von nicht generalisierbaren sozialen Intuitionen durchzogen, hochgradig situativ, abhängig vom Kontext und aufs Engste verbunden mit unserem Selbstwertgefühl und unserem Selbstkonzept. Wiewohl sie sich an Leitplanken orientieren mag, die Philosophen und Gesetzgeber angebracht haben, bleibt sie doch eine emotional komplexe Angelegenheit. Der Gebrauch an Moral lässt sich nicht normieren wie die Größe von Dübeln. Eine Moral ohne subjektive Haltungen, reduziert auf allgemeine Bewertungen des Lebens, bleibt sinnlos. Philosophie der Moral kann deshalb kein Werkzeugkasten sein, um moralische Probleme ein für alle Mal zu lösen. Vielmehr ist sie ein Reflexionsmedium, das moralische Themen in einen Horizont stellt. Man erfindet keine zu Menschen oder Robotern passende Moral, so wie man passende Technik erfindet.

Vor diesem Problem stehen heute die Philosophen, Theologen und Gesellschaftswissenschaftler, die in den vielen Kommissionen und Ethikräten sitzen, um über die Moral des Zweiten Maschinenzeitalters oder die "ethische" Programmierung von KI nachzudenken. Bedroht vom Ziel, klare Analysen und Empfehlungen zu liefern, pflegen sie einen notgedrungen unphilosophischen Umgang mit philosophischen Fragen. Und statt eines ethischen Blicks auf Technik mühen sie sich um einen technischen Blick auf Ethik. Ärgerlich,

ärgerlich, dass man Ethik nicht begründen kann wie einen Wildwasserfluss, den man zum Kanal macht. Moralische Intuition fließt nicht in geregelten Bahnen. Und was kein Programm ist, kann auch nicht programmiert werden.

Bei den in Ethik-Kommissionen zur KI verhandelten Fragen handelt es sich meist nicht um "Probleme", die sich mithilfe der Wissenschaft, einer Technikfolgenabschätzung oder neuerdings "Sozioinformatik" lösen lassen. In der Welt der Moral gibt es keine objektivierenden Aussagen wie in den Naturwissenschaften. Man kann das für belanglos halten; es ist aber gefährlich: Denn je technisierter und instrumenteller man menschliche Moral zerlegt und zergliedert, umso zahnloser wird sie zugleich. Von der ehemaligen Kraft einer philosophisch begründeten Haltung bleibt nur noch ein Gefühl von Einzelproblemen übrig, das nirgendwo einen passenden Ort findet, um sichtbar ein Stoppschild aufzustellen.

Z+

Exklusiv für Abonnenten

Citizen Science

Entwickelt sich mein Kind, wie es sollte?

[<https://www.zeit.de/zeit-wissen/2021/04/citizen-science-kindesentwicklung-eltern-hobbyforschung-buergerwissenschaft>]

Chłodnik

Rote Bete auf Polnisch

[<https://www.zeit.de/zeit-magazin/2021/25/chlodnik-rote-bete-suppe-kalt-rezept-polen-wochenmarkt>]

Christian Thielemann

Titanendämmerung

[<https://www.zeit.de/2021/25/christian-thielemann-dirigent-staatskapelle-dresden-entlassung-semperoper-vertragsverlaengerung>]

Mehr Abotexte → [<https://www.zeit.de/exklusive-zeit-artikel>]

Tatsächlich verlangt die Maschinenethik, Maschinen nicht "ethisch" zu programmieren, also Entscheidungen treffen zu lassen, die über Menschen richten. Ethisch mit Maschinen umzugehen ist genau das Gegenteil davon, sie "ethisch" zu programmieren. Ein Beispiel dafür ist das sogenannte "autonome" Fahren. Voll automatisierte Robo-Cars werden als die Zukunft angesehen. Folglich sollen Ethiker sich Gedanken darüber machen, wie voll automatisiert fahrende Autos in einer Dilemma-Situation entscheiden sollen, wenn sich der Tod von Menschen nicht verhindern lässt. Wer soll überleben und wer

überfahren werden [<https://www.zeit.de/digital/2018-10/autonomes-fahren-kuenstliche-intelligenz-moralisches-dilemma-unfall>]? Kinder oder Rentner? Hunde, Kranke, Straffällige?

Dass das Grundgesetz mit seiner Definition der Menschenwürde solche Überlegungen schon im Ansatz nicht zulässt, hat die deutsche Automobilindustrie bislang zu keinem klaren Bekenntnis gegen ein solches Aufrechnen von Lebenswert genötigt. So steht zu befürchten, dass auch bei uns daran gearbeitet wird, Autos "ethisch" zu programmieren. Da diese Ethik, wie gezeigt, keine ist und zudem dem Grundgesetz widerspricht, müssten sämtliche Ethiker in Deutschland hier laut aufschreien. Stattdessen sitzen in Kommissionen (neben Ausnahmen) hinreichend Fachleute, die leiser kaum sein könnten. Und das, obwohl das zukünftige "autonome" Fahren in unseren Großstädten nicht nur den Lebenswert von Menschen verrechnen soll, sondern auch jeden Verkehrsteilnehmer uneingeschränkt überwachen wird. Robo-Cars in den Innenstädten sind vielleicht die Lösung für unsere Verkehrsprobleme. Viel wahrscheinlicher sind sie nicht einmal das. In jedem Fall aber gehen sie einher mit einem Verstoß gegen das Grundgesetz und einer drastischen Einschränkung unserer Freiheit. Wer das will, will nicht nur einen anderen Verkehr – er will einen anderen Staat.