

# Künstliche Intelligenz

## Hilfreiche Assistenz oder Bedrohung der Menschheit?

TEXT: JOACHIM VON GOTTBURG

Lange Zeit galt die Digitalisierung als hilfreiche Unterstützung des Menschen. Vieles wurde einfacher, schneller - von Schreibprogrammen bis zur digitalen Kommunikation via Smartphone, Sprachassistenten und mehr. Als Erster hat der Google-Chefentwickler Raymond Kurzweil 2005 die These aufgestellt, nach der die künstliche Intelligenz (KI) mit ihren Algorithmen fähiger als die Menschen werden und diese als Krone der Schöpfung ablösen könnte. Informatiker und Philosophen, zuletzt selbst Sam Altman, CEO der Firma OpenAI, die ChatGPT auf den Markt gebracht hat, warnen vor dystopischen Folgen der KI, andere halten das für völlig übertrieben.

\*

**Raymond Kurzweil** geht davon aus, dass sich sowohl biologische als auch technische Entwicklungen exponentiell beschleunigen und die Zeiträume zu neuen, revolutionären Erfindungen immer kürzer werden. Vor 3,8 Mrd. Jahren gab es auf der Erde ausschließlich Einzeller und es dauerte ungefähr 2 Mrd. Jahre, bis auch Mehrzeller entstanden. Diese Entwicklung beschleunigte sich immer weiter: Vom ersten Säugetier bis zum Homo sapiens dauerte es nur 200 Mio. Jahre. Werkzeuge und technische Geräte haben sich zunächst über Jahrhunderte langsam entwickelt, im 20. Jahrhundert hat sich die Technik dann über Autos, Flugzeuge und den Computer erheblich beschleunigt. Inzwischen haben wir Programme, die den Menschen beim *Schach*, aber auch bei komplexeren Spielen wie *Go* schlagen: Werden Programme in Zukunft auch

besser sein als der Mensch, wenn es darum geht, Programme zu schreiben (vgl. Kurzweil 2013)?

Kurzweil vermutet, dass wir bald Supercomputer aus unserer DNA bauen können und Menschen und Technik immer mehr miteinander verschmelzen werden. Noch hat die KI keine eigenen Ziele, keine Motivation und keine Gefühle. Aber: Jede mentale Aktivität spiegelt sich irgendwo in unserem Gehirn, also in unserer Hardware, wider. Bestimmte Neuronen erzeugen bestimmte Botenstoffe. Und das wird man, so Kurzweil, irgendwann nachbauen können.

### Cyborgs: Menschen und Technik verschmelzen

Miniroboter können in unserem Körper Krankheiten erkennen, Viren ausschalten, organische Fehler reparieren und uns früher oder später unsterblich machen. Wir werden immer mehr zu Cyborgs: In dem Film *Matrix* lädt der Held Neo ein Kung-Fu-Programm in seinem Gehirn hoch und beherrscht dadurch die Kampfsportart innerhalb von Sekunden perfekt: So ähnlich werden wir nach Kurzweil in Zukunft lästige Lernprozesse durch digitales Hochladen von Informationen ersetzen und das Lernen beschleunigen und optimieren. Kurzweil prophezeit, dass im Jahr 2045 die KI die Fähigkeiten des Menschen übertreffen wird – er nennt das Singularität: Maschinen können sich selbst optimieren und beschleunigen die technologische Evolution, sodass sie den Menschen nicht mehr braucht.

Auch der Biologe **Paul Rainey** geht davon aus, dass es in der Evolution eine Entwicklung in exponentieller Geschwindigkeit gab, die letztlich immer auf Optimierung angelegt war und ist. Das Zusammenwachsen mit der KI wird, so wie es auch Kurzweil prognostiziert, der nächste Schritt sein: „Ich denke, in gar nicht so ferner Zukunft werden sich Menschen nicht mehr autonom entwickeln. Unsere Spezies ist in Jahrtausenden durch das machtvolle Wirken der natürlichen Auslese entstanden. In jeder Generation konnten sich die am meisten fortpflanzen, deren Nachkommen am besten an ihre Umwelt angepasst sind. So wie das auch bei allen anderen Organismen der Fall ist. Das könnte sich nun ändern. Künftig könnten diejenigen die besten Chancen haben, die gemeinsam mit einer KI am erfolgreichsten agieren. Mensch und KI werden sich in einer Symbiose weiterentwickeln. [...] Die Argumente dafür, dass Menschen und künstliche Intelligenz zusammen eine neue Evolutionseinheit bilden, ergeben sich aus einer Erweiterung der Konzepte, die für die Erklärung der Ursachen von METs (Major Evolutionary Transitions) zentral sind. Menschen und KI könnten im Prinzip eine Evolutionseinheit bilden, wobei sowohl Menschen als auch KI Bestandteile einer einzigen, sich auf höherer Ebene replizierenden Einheit sind“ (Rainey, zitiert nach Herden 2023).

Bereits jetzt sind Menschen mit dem Smartphone verbunden, das etwa die Navigation übernimmt oder über Suchmaschinen unser Denken und unser Gedächtnis unterstützt. „Die Geräte sind nicht das Entscheidende, sondern die Algorithmen, die sich im Laufe unseres Lebens immer besser an unsere Bedürfnisse anpassen. Für eine gemeinsame Evolution wird entscheidend sein, ob wir die Informationen, das, was die Algorithmen auf unseren Geräten gelernt haben, an unsere Kinder weitergeben. Schon jetzt gibt es das Phänomen, dass Eltern ihre alten Handys an ihre Kinder weiterreichen. Und wenn es nicht das Gerät selbst ist, dann sind oft die Apps und Zugangsdaten im Familienbesitz. Bald schon könnten leistungsfähige, angepasste Algorithmen von Vater oder Mutter an die Kinder vererbt werden“ (ebd.).

### Gefahr durch den Wettbewerb

Die Zukunftsforscherin **Amy Webb** sieht die Entwicklung ähnlich wie Kurzweil, allerdings ist sie bezüglich des Verhaltens der KI in der Singularität

weniger optimistisch (Webb 2019). Die größte Gefahr sei, dass es derzeit neun Unternehmen gebe, die künstliche Intelligenz vorantreiben, sechs Unternehmen davon in den USA: Google, Microsoft, Amazon, Facebook, IBM und Apple. Sie müssten Geld verdienen und setzten im Konkurrenzkampf auf Geschwindigkeit statt auf Sorgfalt. Sie nähmen in Kauf, dass die Technik nicht mehr beherrschbar ist. Bei den drei Unternehmen aus China, Baidu, Alibaba und Tencent, gehe es vor allem um die Überwachung von Menschen. Peking will bis 2030 weltweit Vorreiter im Bereich der KI-Innovation sein.

Die Menschheit werde immer abhängiger von künstlicher Intelligenz. Kleine tragbare KI-Geräte könnten uns vor einem drohenden Nährstoffmangel im Körper warnen und die passenden Diäten empfehlen, wir bekämen mikroskopisch kleine KI-Roboter injiziert, die Gebrechen erkennen und ohne jede ärztliche Hilfe heilen könnten. Bei so viel Geschwindigkeit und Wettbewerb seien Fehler vorprogrammiert: Verkehrs- und Gesundheitsnetze könnten lahmgelegt oder der Zugang zum eigenen Kühlschrank gesperrt werden, so Webb weiter (vgl. ebd.).

### Ethische Maßstäbe und gesetzliche Regulierung

Neue KI-Systeme müssten vor der Markteinführung nicht nur auf ihre Funktionsfähigkeit, sondern auch auf ungewollte negative Folgen für die Gesellschaft gründlich getestet werden – das könnten bald neue KI-Systeme in Simulationen erledigen. Allerdings: Selbst wenn die Unternehmen ihre KI-Produkte besser prüfen wollten, wären sie von den kapitalistischen Marktmechanismen gezwungen, sie möglichst schnell auf den Markt zu bringen. Webb hält unter diesen Umständen jede unternehmerische Selbstregulierung für unrealistisch. Deshalb sollten Regierungen verbindliche KI-Richtlinien entwerfen. Nach Webb brauchen wir solide Gesetze und Institutionen, die die Einhaltung dieser ethischen Standards durch die Techfirmen überprüfen. Außerdem sollten beträchtliche Fördergelder in die KI-Branche gepumpt werden, um die Unternehmen vom Druck zu befreien, ihre Produkte – von Gewinnabsichten getrieben – ungeprüft auf den Markt zu bringen.

Webb schlägt vor, mit Verbündeten wie Japan und Kanada eine „Global Alliance of Intelligence“, kurz GAIA, zu gründen, um den Umgang mit der zu erwartenden Explosion künstlicher Intelligenz zu schaffen. In einem solchen internationalen Organ könnten die humanistisch orientierten Regierungen und Unternehmen ihr Wissen teilen, um zusammen Fortschritte zu erzielen, die Rivalen wie China dazu motivieren könnten, sich ihnen anzuschließen, um Wettbewerbsnachteile zu vermeiden.

### Zukunftsvisionen zwischen Dystopie und Optimismus

Der Bonner Philosoph **Markus Gabriel** hat sich mit den Prognosen über die Zukunft der KI beschäftigt: Elon Musk glaube fest an eine dystopische Variante der Singularität. Auch Stephen Hawking habe gegen Ende seines Lebens den Untergang der Menschheit prophezeit und geglaubt, dass die Superintelligenz uns vernichten werde. Amerikanische Milliardäre, aber auch viele Philosophen gingen davon aus, dass die Singularität bevorstehe, etwa der australische Denker David Chalmers und viele seiner Kollegen am Department of Philosophy der New York University (vgl. Gabriel, zitiert nach Wiebicke 2018).

Auch in Deutschland gebe es, so Gabriel, Anhänger der Superintelligenz, etwa den Unternehmer Frank Thelen, der schon vor 2045 damit rechne. Er schreibe selbst Apps, die schon jetzt Superintelligenz-Strukturen aufwiesen. Im Digitalrat der Bundesregierung zeige Andreas Weigend, der in Bonn Philosophie und Physik studiert habe, eine gewisse Neigung in Richtung Singularität (ebd.).

### Was unterscheidet den Menschen von der KI?

Experten wie Robotiker:innen oder Informatiker:innen in Deutschland halten meist wenig von Superintelligenz und Singularität und sehen darin eine kalifornische Verkaufsstrategie, in der aber nicht zwischen Illusion und Wirklichkeit unterschieden werden kann. Nach Gabriel können Menschen über das Denken und über sich selbst reflektieren. Dieser Geist sei für unseren Selektionsvorteil verantwortlich. Die Grundfrage laute: Was ist der Mensch? Das Wesen des Menschen sei so universal, dass über die lokalen, provinziellen Vorstellungen des Menschseins eine Form der Selbstreflexion sowie die Ethik entstehe, die flexibel an Veränderungen anpassbar sei. Eine Maschine könne so etwas nicht leisten. Deshalb glaubt Gabriel nicht an die Theorie der Superintelligenz oder der Singularität.

### Wie intelligent ist die KI?

Die Kulturwissenschaftlerin **Mercedes Bunz** beschreibt im Gespräch mit Richard David Precht KI zwar als intelligent, allerdings anders intelligent als der Mensch. Künstliche Intelligenz klinge sehr nach der des Menschen, daher würde sie den Begriff „Maschinenintelligenz“ vorziehen. KI wisse im Gegensatz zur menschlichen Intelligenz nicht, was sie tun solle, weil ihr das vom Programm vorgegeben werde. Sie werde trainiert und lerne aus vielen Daten, etwas zu verstehen. Inzwischen könne die Maschine zum ersten Mal Bedeutung analysieren, allerdings ausschließlich rechnerisch. Jahrzehntlang hätten Computerwissenschaftler daran gearbeitet, Sprache und Bilder zu verstehen. Aber es sei ein Maschinenverstehen. Bewusstsein hätten die Maschinen nicht, sie besäßen keinen freien Willen und könnten nicht selbst entscheiden, was sie machen wollen. Sie würden programmiert. Allerdings: Die Maschine spreche, als ob sie ein Mensch sei. Erst wenn man die Technik dahinter verstehe, erkenne man, dass die menschliche Sprache imitiert werde. Menschliche Sprache könne man stochastisch gut berechnen, sie sei strukturiert und habe bestimmte grammatische Regeln, die man mathematisch berechnen könne (Bunz, zitiert nach Precht 2023).

### Ethische Überlegungen: Anforderungen an die KI

In Europa gehen die meisten Experten eher von einer schwachen KI aus, die zwar bestimmte Aufgaben schneller und besser erledigen kann als der Mensch, aber ihre Grenzen hat, wie es die Philosophin **Catrin Misselhorn** beschreibt: „In gewissen Bereichen ist die künstliche Intelligenz Menschen ja schon ebenbürtig oder sogar überlegen, etwa im *Schach* oder *Go*. Allerdings hat die menschliche Intelligenz die Besonderheit, dass sie grundsätzlich auf alle möglichen Bereiche anwendbar ist. Eine solche allgemeine Intelligenz ist aus meiner Sicht mit den derzeitigen Methoden der künstlichen Intelligenz nicht zu reproduzieren“ (Misselhorn, zitiert nach Gottberg 2019, S. 26).

Ein Einfühlungsvermögen besitzt die KI aus ihrer Sicht nicht: „Zugespielt formuliert ist emotionale KI etwa so empathisch wie ein Psychopath. Künstliche Systeme können bestenfalls wie ausgebuffte Psychopathen Emotionen erkennen und sozial angemessen darauf reagieren. Diese Fähigkeit dient - ähnlich wie bei den Psychopathen - in vielen Fällen manipulativen Zwecken. Diese Zwecke setzt sich das System aber natürlich nicht selbst, sondern sie sind ihm von seinen Entwicklern vorgegeben. Eigene Ziele können Maschinen nicht entwickeln, weil diese Fähigkeit eng mit dem biologischen Leben verbunden ist. Bewerten und abwägen können Maschinen in einem gewissen Umfang hingegen schon. So kann man ein künstliches System so programmieren, dass es verschiedene Optionen im Hinblick auf eine bestimmte moralische Pflicht bewertet, beispielsweise in der Medizinethik die Pflicht, einem Patienten keinen Schaden zuzufügen. Den unterschiedlichen Handlungsmöglichkeiten

werden dann Zahlenwerte zugewiesen, je nachdem, wie gering oder groß der zu erwartende Schaden ist. Das kann dann auch gegen andere Pflichten abgewogen werden, etwa die Pflicht, zum Wohl des Patienten zu handeln oder seine Autonomie zu respektieren. Was Maschinen jedoch abgeht, ist die menschliche Fähigkeit zur Selbstreflexion und Moralbegründung, die es auf jeder Ebene immer wieder erlaubt, das eigene Denken und Handeln zu hinterfragen“ (ebd.).

Misselhorn schlägt folgende ethische Maßstäbe vor: „Ich habe drei Leitlinien für die Maschinenethik aufgestellt, die auch auf die Ethik der künstlichen Intelligenz im Allgemeinen übertragbar sind: Erstens sollten künstliche Systeme die Selbstbestimmung von Menschen fördern und sie nicht beeinträchtigen. Zweitens sollten sie nicht über Leben und Tod von Menschen entscheiden. Drittens muss sichergestellt werden, dass Menschen stets in einem substanziellen Sinn die Verantwortung übernehmen. An diesen drei Leitlinien können Sie ablesen, dass es um beides geht: um eine Warnung vor den Gefahren, die der ethisch nicht reglementierte Umgang mit bestimmten Technologien beinhalten kann, aber auch darum, wie man die Technologien so gestalten kann, dass sie den menschlichen Bedürfnissen, insbesondere der Selbstbestimmung förderlich sind. Mit diesem letzten Aspekt beschäftigte ich mich etwa in meinen Drittmittelprojekten zur ethischen Bewertung von Assistenzsystemen in der Pflege, der Arbeitswelt und in der Bildung. Hier sehe ich meine Rolle als Beraterin zur ethischen Optimierung der Technologien“ (ebd., S. 23).

### Mehr Sicherheit durch Regulierung?

Es wird immer mehr gefordert, dass die KI nicht losgelöst von gesetzlicher Regulierung weiterentwickelt werden darf. Das EU-Parlament hat am 16. Juni 2023 den AI Act verabschiedet, der jetzt mit dem Rat und den Mitgliedsländern beraten wird und möglichst Ende 2023 in Kraft treten soll. Er sieht drei Risikostufen vor: Die gesellschaftsfeindliche KI, die etwa biometrisch Menschen überwacht und Scores zu ihren sozialen und finanziellen Eigenschaften erhebt, wird verboten. KI der kritischen Infrastruktur muss transparent arbeiten und benötigt eine Zulassung, die regelmäßig neu bewertet wird (mehr dazu siehe Gottberg 2023).

Die Netzaktivistin **Anke Domscheit-Berg** will die Chancen der KI nutzen, aber die Risiken minimieren. Sie kritisiert, dass z.B. die Bundesregierung im letzten Jahr 86-mal KI-Systeme einsetzte, aber nur in einem einzigen Fall eine Bewertung und Evaluierung der Risiken vorgenommen hat. **Timo Greger**, Philosoph und KI-Forscher, glaubt auch nicht an die Vernichtungsszenarien und die Übernahme der Weltherrschaft durch die KI, sondern hält eher die Bewertung ihrer Anwendungen im Kleinen für wichtig, von der Smartwatch über unseren Social-Media-Konsum bis hin zu ChatGPT. Auch **Ralf Otte**, Professor für Industrieautomatisierung und künstliche Intelligenz, hat keine Angst, dass die KI jemals an die zentralen Intelligenzleistungen von Menschen herantreten wird. Die Gefahr geht nach seiner Auffassung nicht von der Technik aus, sondern von den Menschen, die diese Technik einsetzen.

### Warnung vor KI als Marketing?

Im März 2023 haben in einem vom Center for AI Safety (Zentrum für KI-Sicherheit) in San Francisco veröffentlichten Brief mehr als 350 Fachleute sinngemäß gewarnt: „Es sollte global priorisiert werden, das Risiko der Auslöschung durch KI zu verringern - auf einer Stufe mit anderen Risiken für die gesamte Gesellschaft, wie etwa Pandemien und Nuklearkrieg.“ Dass künstliche Intelligenz die Menschheit gefährdet, wird allerdings von den meisten deutschen Experten nicht angenommen. Anke Domscheit-Berg hat keine Angst davor, dass noch in ihrer Lebenszeit „irgendeine KI die Machete schwingt und uns alle um die Ecke bringt“ (Domscheit-Berg, zitiert nach Maull 2023). Die

Gefahren durch die KI seien tatsächlich bereits da, sie seien vielfältig und auch groß und gefährlich. Aber den offenen Brief, den auch Sam Altman, CEO der Firma OpenAI, die ChatGPT betreibt, unterschrieben hat, hält Domscheit-Berg wie viele andere auch für ein gigantisches Ablenkungsmanöver. Insgesamt steuere die gesellschaftliche Debatte in zwei unterschiedlich große Lager, „wo der eine glaubt, die bringt uns um, und das andere Lager, die bringt uns nicht um“ (ebd.).

Als konkrete Gefahr werden die neuen Manipulationsmöglichkeiten gesehen, beispielsweise könnte ein nicht mehr verifizierbares Video hergestellt werden, in dem Olaf Scholz Russland den Krieg erkläre. Auch könnten Menschen diskriminiert werden, weil beispielsweise automatisierte Personalrekrutierungsprogramme mit Daten trainiert worden seien, die aus einer nicht gleichberechtigten Welt stammten. Wie bereits in den sozialen Medien würde sich die Macht immer mehr in den gleichen Händen zentralisieren, bei Microsoft, Google und den übrigen Big Companys. Davon lenke Sam Altman im Moment leider relativ erfolgreich ab (ebd.).

### Vernichtung von Arbeitsplätzen?

KI, so befürchten viele, könnte Arbeitsplätze kosten. Sie kann immer besser Texte und Sprache übersetzen und viele Dolmetscher:innen machen sich Sorgen, durch KI ersetzt zu werden. Das trifft vor allem eher geisteswissenschaftlich orientierte Berufe, doch selbst Abteilungen von Hedgefonds-Managern sind bei einer indischen Bank entlassen worden. In den USA sollen bei 80 % der aktuellen Jobs etwa 10 bis 20 % der Tätigkeiten durch KI potenziell ersetzbar sein. Andere Bereiche, etwa das Handwerk, werden dagegen wohl weniger betroffen sein.

### Fazit

Die Einschätzung der Gefahren von KI hängt vor allem mit der Frage zusammen: Kann eine Maschine aus sich heraus eine Absicht und eine Motivation entwickeln? Kann sie einen eigenen Willen besitzen, etwa den der Vernichtung oder Unterdrückung der Menschen? Oder folgt die Maschine immer nur der Absicht und der Ethik ihrer Programme, die aber von Menschen geschaffen werden und zu deren Unterstützung dienen. Dann würden die Gefahren weniger von der KI ausgehen, sondern von denen, die sie gemäß ihren konstruktiven oder destruktiven Absichten programmieren. Wie sich das entwickeln wird, kann nicht mit Sicherheit prognostiziert werden: Aber Vorsicht ist geboten.

\*

#### **Literatur:**

**Gottberg, J. von:** *Programmierte Ethik. Künstliche Intelligenz ohne Entscheidungsspielräume.* Ein Gespräch mit Catrin Misselhorn. In: tv diskurs, Ausgabe 90, 4/2019/23, S. 22-27. Abrufbar unter: <https://mediendiskurs.online>

**Gottberg, J. von:** *Die Geister, die ich rief. KI-Entwickler warnen vor den Gefahren ihrer Produkte.* In: mediendiskurs, 15.06.2023. Abrufbar unter: <https://mediendiskurs.online>

**Herden, B.:** *Künstliche Intelligenz: „In Zukunft werden sich Menschen nicht mehr autonom entwickeln“.* Gespräch mit Paul Rainey. In: Tagesspiegel, 30.03.2023. Abrufbar unter: <https://www.tagesspiegel.de>

**Kurzweil, R.:** *Menschheit 2.0. Die Singularität naht.* Berlin 2013

**Mauil, D.:** *Vernichtung durch KI - Wie gefährlich ist künstliche Intelligenz?.* Ein Gespräch mit Anke Domscheit-Berg, Timo Greger und Ralf Otte. In: SWR 2 Forum, 13.06.2023. Abrufbar unter: <https://www.swr.de>

**Precht, R. D.:** *Kein Platz mehr für den Menschen - Macht KI uns überflüssig? Ein Gespräch mit Mercedes Bunz.* In: ZDF-Mediathek, 04.06.2023. Abrufbar unter: <https://www.zdf.de>

**Webb, A.:** *Die großen Neun. Wie wir die Tech-Titanen bändigen und eine künstliche Intelligenz zum Wohle aller entwickeln können.* Kulmbach 2019

**Wiebicke, J.:** *Philosoph Markus Gabriel über den Menschen und KI.* In: WDR 5: Das philosophische Radio, 13.09.2018. Abrufbar unter: <https://www.youtube.com>



Prof. Joachim von Gottberg ist Chefredakteur von *mediendiskurs*.