

# Künstliche Intelligenz

## Häufig gestellte Fragen

### Einleitung

Ob als kluge Assistentin auf dem Smartphone, als clevere Routenführung im Auto oder als smarte Unterstützung bei der medizinischen Diagnose – überall wirken lernende Algorithmen im Hintergrund. Unter der Bezeichnung „Künstliche Intelligenz“ (KI) kommen diese digitalen Technologien in immer mehr Bereichen zur Anwendung. Sie stehen in der Forschung, der Wirtschaft und auch der Politik zurzeit ganz oben auf der Agenda.

Durch die Verfügbarkeit von riesigen Datenmengen, den Zugang zu einer für ihre Verarbeitung notwendigen Rechenleistung sowie Algorithmen, die wie das menschliche Gehirn aufgebaut sind, können IT-Systeme jetzt über Fähigkeiten wie situatives Wahrnehmen, Kommunizieren, Planen, Handeln, Schlussfolgern oder Lernen verfügen. Sie können damit zum Teil Aufgaben schneller und präziser bewältigen als der Mensch. So ist das Auffinden von Mustern in einer großen Menge von Informationen mit Hilfe einer KI leicht zu lösen. Aber ist sie deshalb schon intelligent?

Ohne Fachkenntnisse ist die Funktionsweise einer KI nur noch schwer zu verstehen. Auch ihre Auswirkungen auf Wirtschaft, Wissenschaft, Gesellschaft und Kultur können dann kaum realistisch eingeschätzt werden. An ihre Stelle treten oft übertriebene Erwartungen („Superintelligenz“) oder überzogene Befürchtungen („Roboterherrschaft“). Mit unserem Infoplip wollen wir die häufig gestellten Fragen zu den technischen Grundlagen und Begriffen der Künstlichen Intelligenz beantworten, an Beispielen die Anwendungsmöglichkeiten erklären und die gesellschaftspolitischen sowie ethischen Problemfelder aufzeigen.

### Tobias Wangermann

Konrad-Adenauer-Stiftung

### Was ist Künstliche Intelligenz?

#### Wie funktioniert maschinelles Lernen?

#### Welche Bedeutung haben Daten für die KI?

Diese und andere Fragen zur Künstlichen Intelligenz finden Sie in dieser Broschüre einfach und verständlich beantwortet. Wenn Sie mehr Informationen zum Thema suchen, dann gehen Sie im Internet auf unsere Themenseite [www.kas.de/ki](http://www.kas.de/ki) oder auf unser aktuelles Portal zu Künstlicher Intelligenz und Robotik [www.meschmasch.de](http://www.meschmasch.de).

Im Auftrag der Konrad-Adenauer-Stiftung  
herausgegeben von Tobias Wangermann

### Erarbeitet von

Dr. Norbert Arnold (Konrad-Adenauer-Stiftung),  
Dr. Aljoscha Burchardt (DFKI – Deutsches Forschungszentrum Künstliche Intelligenz),  
Stefan Gehrke (bfnd – buero fuer neues denken),  
Philipp Otto/Eike Gräf (iRights.lab) und  
Dr. Ilja Radosch (Fraunhofer FOKUS).

**Design und Layout:** yellow too Pasiiek Horntrich GbR

**Druck:** Infoplip Medien GmbH

Konrad-Adenauer-Stiftung e. V. 2018, Sankt Augustin/Berlin



Die Texte dieses Werkes sind lizenziert unter den Bedingungen von „Creative Commons Namensnennung-Weitergabe unter gleichen Bedingungen 4.0“, CC BY-SA 4.0 (abrufbar unter <https://creativecommons.org/licenses/by-sa/4.0/legalcode.de>).

ISBN 978-3-95721-467-6

Bereits der Philosoph Gottfried Wilhelm Leibniz (1646–1716) träumte von einer Maschine, die Menschen helfen würde, ihre Aussagen mithilfe von logischen Schlüssen auf Richtigkeit zu prüfen. Er wollte uns damit leidige Diskussionen ersparen, die schlicht auf falschen Tatsachen oder Annahmen beruhen – ein hehrer Gedanke.

In den 1950er-Jahren kam die Idee auf, die ersten Computer („Rechner“) zu nutzen, um verschiedene geistige Fähigkeiten zu simulieren. Die Künstliche Intelligenz (KI) war geboren. Da es im Kalten Krieg an Übersetzern mangelte, die aus dem Russischen ins Englische übersetzen konnten, war eine der ersten Forschungsrichtungen die Maschinelle Übersetzung. Auf die Frage, wie lange es dauern würde, bis Computer Texte übersetzen könnten, antworteten die Wissenschaftler etwas fahrlässig: „Drei bis fünf Jahre“. Heute wissen wir, dass 60 Jahre die bessere Antwort gewesen wäre. Erst mit dem Umstieg auf neuronale Netze (→ [Was sind neuronale Netzwerke?](#)) konnten Übersetzungsprogramme endlich eine sehr hohe Qualität erreichen.

Texte übersetzen, Bilder im Internet suchen, Roboter steuern – das sind nur einige Anwendungen von KI, die Einzug in unseren Alltag und unser Arbeitsleben halten. Ihnen gemeinsam sind die eng umrissenen Einsatzgebiete. Im Gegensatz hierzu können wir Menschen uns in kürzester Zeit auf neue Gegebenheiten einstellen und Aufgaben meistern, die wir vorher nicht kannten. Die KI-Anwendungen müssen jedoch für jede einzelne Anwendung sehr aufwändig angepasst werden, oft durch maschinelles Lernen (→ [Wie funktioniert maschinelles Lernen?](#)). Man spricht hier von schwacher KI im Gegensatz zu starker KI (→ [Was ist schwache und was starke KI?](#)).

Eine allgemeingültige Definition von KI kann es nicht geben, da menschliche Intelligenz schon nicht eindeutig definiert ist. Man denke nur an Begriffe wie Handlungsintelligenz („Bauernschläue“), kognitive Intelligenz, emotionale Intelligenz, soziale Intelligenz, etc. Diese lassen sich nicht auf einen Nenner bringen.

Als Arbeitsdefinition für KI kann das Folgende dienen: Künstliche Intelligenz ist die Eigenschaft eines IT-Systems, der menschlichen Kognition ähnliche Fähigkeiten zu zeigen. Die einzelne Fähigkeit kann ansatzweise erkennbar sein, wie etwa die Dialogfähigkeit von Smartphones. Sie kann aber auch über die menschliche Leistungsfähigkeit hinausgehen, wie bei der Auswertung zehntausender MRT-Scans. KI-Systeme verfügen in unterschiedlichen Anteilen über bestimmte Kernfähigkeiten wie situatives Wahrnehmen, Kommunizieren, Planen, Handeln, Schlussfolgern oder Lernen.

**Dr. Aljoscha Burchardt**  
DFKI

Alle kommerziellen Anwendungen von KI und das Gros der Wissenschaftler zielen auf die oben erwähnten Einzelfähigkeiten ab (→ [Was ist künstliche Intelligenz?](#)). Dies bezeichnet man als schwache KI. Die starke KI zielt auf eine allgemeine KI ab, was im Englischen oft als *general artificial intelligence* bezeichnet wird. Letztlich wäre das ein Homunculus, ein Nachbau des Menschen, der genauso flexibel auf Situationen reagieren kann wie ein Mensch, der selbst lernen, sich bewegen und kommunizieren kann wie wir.

Die starke KI hat seit der Geschichte von Frankenstein's Sohn die Menschen beflügelt und ist das Thema vieler Hollywood-Filme – oft mit unangenehmem Ausgang für das Menschengeschlecht.

Aus eher philosophischer Sicht gibt es seit langem Debatten darüber, wie genau man erkennen kann, ob man es mit einer starken KI oder mit einem Menschen zu tun hat. Einer der ältesten Tests ist der Turing-Test, den der KI-Pionier Alan Turing bereits um 1950 entwarf. Kurz gesagt, wird hierbei eine Interview-Situation angenommen, in der in einer Kiste eine Maschine steckt und in einer ein Mensch. Wenn ein Interviewer nach einer bestimmten Zeit nicht sagen kann, in welcher Kiste wer steckt, hat die Maschine den Test bestanden. Der Test ist an die Vorstellungskraft seiner Zeit gebunden und berücksichtigt vorwiegend den Aspekt der Sprach- und Dialogfähigkeit, wobei das Interview über Tastatur und

Bildschirm erfolgt.

Unter Wissenschaftlern ist umstritten, ob und wann eine starke KI in der Zukunft konstruiert werden kann. Bisher sind alle Versuche gescheitert, das unendliche Wissen, über das Menschen verfügen, auch nur in Bruchteilen auf die Maschinen zu bringen – angefangen von relativ überschaubaren Fakten wie den Naturgesetzen (z. B. „alles fällt nach unten“) bis hin zu den komplizierteren Denkmustern („Weiß die Frau, dass ich nicht weiß, ob sie mich kennt?“). Ohne immenses Vorwissen ist auf der anderen Seite aber selbständiges Lernen nicht vorstellbar. (→ [Wie funktioniert maschinelles Lernen?](#)) Die Tatsache, dass unsere Menschenkinder, grob gesagt, das erste Viertel ihres Lebens in der Ausbildung sind, zeigt, wie schwierig und langwierig es ist. Derzeitige Methoden geben nicht einmal eine theoretische Möglichkeit für starke KI her.

Bei der schwachen KI steht der Mensch im Mittelpunkt und gibt die Ziele vor. Die Technik ist Werkzeug und Hilfsmittel, kann aber für den Menschen von zentraler Bedeutung sein, etwa bei einem Rollstuhl, der sich über Gesten oder direkt über Hirnströme steuern lässt.

**Dr. Aljoscha Burchardt**  
DFKI

Das maschinelle Lernen ist derzeit die erfolgreichste Art, Computern ein wenig Intelligenz einzuhauchen. Die anderen beiden Arten sind zunächst der Algorithmus, bei dem, wie in einem Kochrezept, der Maschine alle zu erledigenden Schritte genau vorgegeben sind. Da die wenigsten Aufgaben in unserer Welt von strikt algorithmischer Natur sind, kommt diese Methode nur in speziellen Fällen zum Einsatz, etwa beim Taschenrechner.

Die zweite Möglichkeit wäre, einen kleinen Ausschnitt von unserem Wissen auf dem Computer zu modellieren. Man spricht hier auch von Expertensystemen, die oft mit Wenn-Dann-Regeln funktionieren. Als Beispiel können Smarte Assistenten dienen. Diese können bei bestimmten Schlüsselwörtern wie „Flug buchen“ einigermaßen flexibel durch verschiedene Dialogbausteine vom Nutzer abfragen, wann und wohin er fliegen möchte. Die Modellierung des Wissens ist sehr aufwändig. Daher ist der Funktionsumfang von Expertensystemen eingeschränkt. Ihr Vorteil liegt in der Nachvollziehbarkeit des Systemverhaltens.

Beim maschinellen Lernen geht man einen dritten Weg. Hier überlässt man es mathematisch-statistischen Methoden, aus großen Datenmengen selbst Muster zu „lernen“ oder Handlungen abzuleiten. Zum Beispiel gibt man einem System 1.000 Bilder von Katzen und 1.000 Bilder von Kanarienvögeln mit der entsprechenden Bezeichnung als Zielvorgabe. Neue Bilder können dann automatisch als Katze oder Kanarienvogel klassifiziert werden. (→ [Wie funktioniert Bilderkennung mit KI?](#)) Dabei hat man wenig Kontrolle darüber, welche der von den Ingenieuren vorgegebenen Eigenschaften der Daten zum Lernen herangezogen werden, wie z. B. Farbe oder Größenverhältnisse. Daher sollte man möglichst ausgewogene und umfassende Trainingsdaten nutzen, die ähnlich den Daten sind, die später klassifiziert werden. (→ [Welche Bedeutung haben Daten für die KI?](#)) Zeigt man dem beschriebenen System etwa ein Bild einer Banane, wird es sie vermutlich als Kanarienvogel klassifizieren, dem besten Treffer. Das System hat kein Wissen über Merkmale wie Federn oder Schnurrhaare. Es führt reines statistisches Muster-Erkennen auf den Pixeln der Bilder aus. Gerade das macht diese Systeme aber flexibel und relativ einfach in der Erzeugung. Nachteil ist die schlechte Nachvollziehbarkeit der Ergebnisse und damit einhergehend das Fehlen von Leistungsgarantien.

Verschiedene Arten des maschinellen Lernens, unterscheiden sich z. B. darin, ob die Trainingsdaten von Menschen ausgezeichnet sein müssen oder nicht, ob ein System auf Basis von menschlicher Bewertung der Ergebnisse dauerhaft lernt oder ob es einen automatischen Test gibt.

**Dr. Aljoscha Burchardt**  
DFKI

Eine derzeit besonders erfolgreiche Art des maschinellen Lernens (→ [Wie funktioniert maschinelles Lernen?](#)) sind neuronale Netze. Diese sind mathematische Gebilde, welche der Struktur bestimmter Teile unseres Gehirns nachempfunden sind. Unser Gehirn besteht aus einer riesigen Anzahl miteinander vernetzter Nervenzellen. Die Mächtigkeit unseres Gehirns resultiert aus der massiven parallelen Verarbeitung von externen Reizen sowie der Flexibilität der Verschaltung. So können etwa nach einem Schlaganfall nicht betroffene Areale des Hirns die Funktionen des zerstörten Gewebes übernehmen. Wie genau in unserem Gehirn aus den an sich einfachen Bausteinen komplexe Erkennens- und Denkprozesse entstehen, ist aber weitgehend unbekannt.

Die künstlichen neuronale Netze sind die mathematische Simulation von einer Anzahl von nachempfundenen und vereinfachten Neuronen, die in Schichten miteinander verbunden sind. Mehr haben sie mit dem menschlichen Denken nicht gemeinsam. Während die mathematischen Grundlagen schon über 50 Jahre bekannt sind, ist es erst vor kurzem möglich geworden, durch verbesserte Rechenleistung, neuartige Algorithmen und große Mengen vorhandener Daten, neuronale Netze für bestimmte Aufgaben wie das Klassifizieren von Bildern oder das Übersetzen von Texten zu trainieren.

Während des Trainings verändern sich die Gewichtungen der Verbindungen zwischen den Neuronen, die die Reizverarbeitung im Netz steuern. Der oft beschriebene „Black-Box-Charakter“ der Netze rührt daher, dass man heute nicht in der Lage ist, nachzuvollziehen, was die Netze dabei eigentlich wirklich lernen (vgl. das menschliche Gehirn). Während das beim traditionellen maschinellen Lernen vorher auch schon in vielen Fällen galt, so gilt es bei den Netzen im Besonderen, da hier nur die Eingabe und gewünschte Ausgabe vorgegeben werden und die Systeme alles Weitere „end-to-end“ selbst erledigen. Am Ende steht ein auf die zu lösende Aufgabe optimiertes Netz mit den dafür erforderlichen gewichteten Verbindungen zwischen den Neuronen – eine Interpretation des Gelernten ist so nicht mehr möglich.

Praktisch alle jüngsten Erfolge der KI vom Go-Spielen bis zum autonomen Fahrzeug basieren in Teilen auf maschinellem Lernen mit neuronalen Netzen. In der Forschung arbeitet man intensiv daran, die Erklärbarkeit zu verbessern und verschiedene bisher getrennte Aufgaben wie das Bild- und Textverarbeiten zusammenzubringen.

**Dr. Aljoscha Burchardt**  
DFKI

Als „technologische Singularität“ wird der Zeitpunkt bezeichnet, an dem KI leistungsfähiger als die menschliche Intelligenz ist. Eine solche „Superintelligenz“ gibt es noch nicht, und es ist nicht erkennbar, dass sie in naher Zukunft entwickelt werden könnte.

„Singularität“ und „Künstliche Superintelligenz“ sind Begriffe, die eher technologischen Utopien als realer Wissenschaft und Forschung zuzuordnen sind. Sie verweisen auf eine ferne Zukunft. Trotz des utopisch-visionären Charakters kann aber grundsätzlich nicht ausgeschlossen werden, dass eine solche Superintelligenz irgendwann einmal entwickelt wird und damit eine Singularität eintritt.

Obwohl eine technologische Singularität unrealistisch ist, ist sie doch immer wieder Thema aktueller Diskussionen. Es gibt Stimmen, die vor Gefahren für die Menschheit warnen: KI, die dem Menschen überlegen sei, wende sich gegen sie, unterjochte sie oder rotte sie gar aus. Etwa aufgrund einer fehlerhaften Programmierung, die die Menschen als minderwertig ansehe, oder, weil KI Eigeninteressen entwickle und sich „bewusst“ gegen die Menschheit wende, um die Herrschaft über die ganze Welt zu übernehmen.

Die aktuelle KI-Forschung ist jedoch von der Entwicklung einer Superintelligenz weit entfernt. Die Leistungsfähigkeit von KI ist zwar heute schon sehr beeindruckend: Watson, DeepBlue und AlphaGo zeigen, was KI alles kann. Ihre Intelligenz ist jedoch immer noch sehr speziell: Entweder spielt intelligente Software sehr gut Schach oder Go, brilliert bei *Jeopardy!*, erkennt Sprache oder identifiziert Gesichter. Doch es ist stets spezialisierte KI, eingegrenzt auf enge Anwendungsfelder. Keine intelligente Software kann alles gleichzeitig, Menschen sehr wohl. Sie haben eine allgemeine Intelligenz von hohem Leistungsvermögen. Eine „allgemeine KI“ gibt es nicht. (→ *Was ist „schwache“ und was „starke“ KI?*) Und nach Aussagen von Wissenschaftlern ist die KI-Forschung meilenweit davon entfernt. Könnte sie eine allgemeine KI entwickeln, wäre es immer noch keine Superintelligenz, die besser ist als der Mensch. Eine Superintelligenz und damit das Eintreten einer Singularität liegen also in weiter Ferne. Aus heutiger Perspektive sind sie mehr Science Fiction als Science.

**Dr. Norbert Arnold**  
Konrad-Adenauer-Stiftung

### Lesetipps:

Nick Bostrom: **Superintelligenz. Szenarien einer kommenden Revolution.** Suhrkamp-Verlag, Berlin 2014.

Klaus Mainzer: **Künstliche Intelligenz – Wann übernehmen die Maschinen?** Springer-Verlag, Berlin/Heidelberg 2016.

Die Frage, ob KI ein Bewusstsein entwickeln kann, lässt sich nicht eindeutig beantworten. Zunächst muss geklärt werden, was überhaupt unter „Bewusstsein“ verstanden wird. Unterschiedliche Eigenschaften werden mit diesem Begriff verbunden. Eine einfache Form liegt vor, wenn Organismen nicht nur auf äußere Reize reagieren, sondern sie als inneres Phänomen erleben. Diese Art Bewusstsein haben Tiere mit differenzierten Gehirnen. Eine höhere Bewusstseinsstufe liegt vor, wenn Lebewesen Gedanken, Ziele und Interessen entwickeln. Und schließlich gibt es Lebewesen – der Mensch, wahrscheinlich Primaten und andere hochentwickelte Tiere –, die „selbst-bewusst“ sind, sich als existierend und als Individuum wahrnehmen.

Der philosophische Dualismus geht davon aus, dass Bewusstsein nicht aus physikalischen Prozessen abgeleitet werden kann, sondern von grundsätzlich anderer – geistiger – Qualität ist. Materialistische Positionen nehmen an, dass Bewusstsein eine Funktion des Gehirns ist und dass neurologische Vorgänge mit Bewusstsein korrelieren. Dafür sprechen empirische Fakten, z. B.: Wenn Teile des Gehirns nicht funktionieren, führt dies zu Veränderungen des Bewusstseins. Das menschliche Gehirn besteht aus 100 Mrd. Neuronen, die hochgradig vernetzt sind. Eine Hypothese besagt, dass Bewusstsein entsteht, sobald diese neuronale Vernetzung eine hohe Komplexität erreicht hat, so dass nicht mehr nur externe Reize verarbeitet werden, sondern zusätzlich

auch gehirn-interne Signale. Dies hat weitreichende Folgen für KI. Denn dort werden künstliche neuronale Netze (KNN, lernfähige Computerprogramme, die natürliche neuronaler Netze simulieren) verwendet. (→ *Was sind neuronale Netzwerke?*) Sind KNN ausreichend komplex, dann könnte, dieser Hypothese folgend, künstliches Bewusstsein entstehen. So wie die Neurowissenschaften annehmen, dass Bewusstsein eine Funktion des Gehirns ist, geht KI-Forschung davon aus, dass künstliches Bewusstsein als Folge komplexer KNN entsteht.

Ob dies wirklich zutrifft, ist derzeit offen. Noch haben die KNN die notwendige Komplexität bei weitem nicht erreicht. Hinzu kommen theoretische Einwände, die darauf hinweisen, dass es ein Unterschied sei, Bewusstsein zu „haben“ (wie der Mensch) und Bewusstsein nur zu „simulieren“ (wie KNN/KI). Spannend ist die Frage, welche Tiefe künstliches Bewusstsein entwickeln könnte. Entwickelt KI irgendwann einmal sogar eigene Interessen?

**Dr. Norbert Arnold**  
Konrad-Adenauer-Stiftung

### Lesetipp:

Ray Kurzweil: **How to create a mind.** Penguin Books, New York, 2012.

Das Erkennen von Bildern gehört zu den größten Herausforderungen der Künstlichen Intelligenz. Während man Stimmen oder Musik in Wellen darstellen kann und sie damit quasi unverwechselbar und wiedererkennbar werden, funktioniert dies bei Bildern nicht so einfach.

Sehr gut ausprobieren kann das jeder beispielsweise mit der nicht so bekannten umgekehrten Bildersuche von Google (<https://images.google.de/>). Auf der Webseite lässt sich ein Bild hochladen oder eine Bild-Adresse eingeben. Google sucht dann nach vergleichbaren Bildern. Gelingt dies z. B. mit einem Foto des Eiffelturms oder anderen millionenfach abgebildeten Objekten noch sehr gut, kommt die Suchmaschine bei Fotos von Gegenständen, Tieren oder gar Menschen (noch) sehr schnell an ihre Grenzen.

Was muss eine Software können, um Bilder zu erkennen, oder besser, um sie zu verstehen? Für einen Computer ist ein Bild zunächst nicht mehr als eine Ansammlung von Millionen von Pixeln, also Bildpunkten. Erst die Gesamtheit dieser Punkte, ihre Farbe, ihr Kontrast, ihre Helligkeit, formen im Auge des Betrachters ein Bild. Die Software sucht nun in dieser Anordnung von Punkten eine Ähnlichkeit zu gespeicherten Bildern. Voraussetzung ist also zunächst eine Datenbank mit einer großen Anzahl von Bildern; je mehr, desto besser. (→ [Welche Bedeutung haben Daten für die KI?](#)) Nun sind Bilder aber aus unterschiedlichen Perspektiven vor unterschiedlichen Hintergründen aufgenommen worden. Alles Dinge, die vom Wesentlichen, z. B. einem Gesicht, ablenken.

So sollen z. B. die neuen Regeln für Passfotos (heller Hintergrund, keine Brille, keine Emotionen) dafür sorgen, dass die Bilder von Maschinen gelesen und erkannt werden können. Bei der Gesichtserkennung spielen zudem noch der Abstand von z. B. Augen, Nase, Mund und anderen besonderen Merkmalen eine Rolle. Diese werden in einer Vektoren-Matrix erfasst und damit vergleichbar gemacht. Aber auch das lässt sich austricksen.

Viele meinen gar, dass eine Künstliche Intelligenz, die den Namen verdient, zumindest bei der Bilderkennung noch nicht in Sicht ist. Dazu fehle die wesentliche Voraussetzung, die bisher nur Menschen einigermaßen fehlerfrei beherrschen: ein gewisses Grundverständnis der Dinge und ihrer Beziehungen zueinander.

**Stefan Gehrke**  
bfnd

### Lesetipp:

Google-Forscher haben einen Farbfleck entwickelt, der die automatische Bilderkennung des Computers außer Kraft setzt. Ist dieser knallbunte, gemusterte Fleck im Bild, sieht die Software überall nur noch Toaster. Alle anderen Bildinhalte werden ignoriert. (Quelle: <https://arxiv.org/pdf/1712.09665.pdf>, aufgerufen am 16.08.2018)

Filme des Genres Science Fiction sind ihrer Zukunft oft um Jahrzehnte voraus. Und sie haben den Vorteil, dass nichts in ihnen der Realität entsprechen muss. So ist es auch bei sprechenden Maschinen: HAL 9000, der mächtige Computer im Film 2001: *Odyssee im Weltraum* aus dem Jahr 1968, kommuniziert unentwegt mit der Crew des Raumschiffs und entwickelt sogar menschliche Züge. Ab 1982 plauderte dann das Wunderauto K.I.T.T. aus der TV-Serie *Knight Rider* mit seinem Fahrer Michael Knight.

Was damals die Kino- und Fernsehzuschauer begeistert hat, ist in der Realität erst seit Kurzem möglich. Inzwischen stehen Amazons *Alexa*, Googles *Assistant*, Microsofts *Cortana*, Apples *Siri* und Samsungs *S-Voice* im Wettbewerb um die Aufmerksamkeit der Smartphone-Nutzer.

Dabei sind auch bei der Spracherkennung Daten, Algorithmen und Rechenleistung die Basis für ein funktionierendes System. Sprache muss dabei – ähnlich wie Bilder in Pixel – so zerlegt werden, dass ein Computer sie versteht.

(→ [Wie funktioniert Bilderkennung mit KI?](#)) Sprache ist eine Folge von grundlegenden Toneinheiten, die von der menschlichen Sprechereinheit im Hals erzeugt werden. Gesprochene Wörter bestehen aus aufeinanderfolgenden Toneinheiten. Um ein Wort zu verstehen, müssen wir versuchen, die ihm zugrunde liegenden Toneinheiten zu identifizieren. Beim Menschen erfolgt dieser Vorgang im Innenohr und wird dann

im Gehirn mit einer Bedeutung verbunden.

Was für uns selbstverständlich erscheint, ist für den Computer eine große Herausforderung. So muss er möglichst viele akustische Modelle einer Sprache gespeichert haben. Das heißt, er muss den Klang von möglichst vielen Wörtern kennen, damit er während eines Gesprächs diese Modelle mit dem Gehörten vergleichen und entscheiden kann, welches Wort wahrscheinlich gemeint war. Mathematisch bezeichnet man diesen Vorgang als Kreuzkorrelation.

Damit der Computer diese Sprachmodelle erstellen kann, benötigt er zunächst eine enorme Menge an Sprachdaten mit denen er das „Gehörte“ vergleichen kann. Je mehr Daten zur Verfügung stehen, desto genauer ist die Spracherkennung.

In der Praxis funktioniert das so: Man spricht in das Smartphone, die Spracherkennung verändert das analoge Signal der Stimme, das am Mikrofon ankommt, in ein digitales. Dieses Signal wird in Frequenzen umgewandelt, das Schallbild (Spektrogramm) der Stimme anschließend in kleine Stücken zerlegt und an verschiedene Server weltweit versendet. Die Software verarbeitet diese Datenpakete, indem es in einer Schicht einzelne Elemente identifiziert (Vokale und Konsonanten). Anschließend erkennt eine weitere Schicht die Gruppierungen der einzelnen Laute und setzt diese Schritt für Schritt weiter zusammen bis ein Wort entsteht.

**Stefan Gehrke**  
bfnd

Wer heute an Roboter in der Industrie denkt, hat meistens die einarmigen Maschinen der Automobilhersteller vor Augen, die unermüdlich, Tag und Nacht, präzise perfekte Schweißnähte setzen, Schrauben mit dem exakten Drehmoment anziehen oder Windschutzscheiben passgenau einsetzen.

Eine Arbeit, die der Mensch in solcher Perfektion und Präzision nicht dauerhaft ausüben kann. Diese Tätigkeit der Maschinen setzt kluge Programmierung und perfekte Platzierung der Werkstücke voraus, damit der Roboter seine automatisierten Abläufe durchführen kann. Im Heimbereich denken wir vielleicht an Saugroboter oder autonome Rasenmäher, die Hindernissen wie Blumen oder Katzen ausweichen. Auch wenn ohne die Einführung von Robotern, insbesondere in der Automobilproduktion, eine wettbewerbsfähige Industrie in Deutschland heute nicht mehr existieren würde – mit Intelligenz hat diese Arbeit noch recht wenig zu tun.

Was passiert aber nun, wenn diese Automatisierungstechnologie vereint wird mit Künstlicher Intelligenz? Kombiniert mit Computerprogrammen, die aus einer Menge von Fakten und Regeln korrekte Schlüsse ziehen können. Dazu kommt der technische Fortschritt, der es ermöglicht, dass immer kleinere und leistungsfähigere Kameras, Sensoren und Computerchips entwickelt werden. So entstehen Maschinen von hoher struktureller Komplexität mit einer enormen algorithmischen Kapazität.

Diese Maschinen können dann nicht nur eine Mutter mit exakt definiertem Drehmoment auf eine Schraube drehen, sondern sie erkennen auch, wo diese Schraube platziert werden muss, welche Schraube es sein muss, welches Werkzeug dafür nötig ist und wann diese Schraube gedreht werden muss.

Wenn die technischen Hindernisse wie Systemkomplexität und Dateninkompatibilität überwunden werden, entstehen neue Dienstleistungsmodelle. Diese basieren auf Echtzeitdaten der Roboter, die von Sensoren erfasst werden. Zudem entwickelt sich der Markt für Cloud-Robotik: Hier lassen sich Daten eines Roboters mit Daten von anderen Robotern synchronisieren und miteinander vergleichen – an demselben oder an unterschiedlichen Standorten. Das Cloud-Netzwerk erlaubt es allen miteinander vernetzten Robotern, die gleichen Aktivitäten auszuführen. Damit werden Parameter wie Geschwindigkeit, Winkel oder Krafteinsatz des Roboters optimiert.

Welche Auswirkungen diese Entwicklung auf die Zukunft der Arbeit und die Art wie wir künftig produzieren hat, ob der Mensch in der Produktion überflüssig wird oder ob er nur einen „klugen“ Robotergehilfen zur Seite bekommt, wird die Zukunft zeigen.

**Stefan Gehrke**  
bfnd

Die Visionen für die Zukunft des automatisierten Verkehrs sind verlockend. Die Künstliche Intelligenz ist dabei vor allem in drei Bereichen von großer Bedeutung: für die Wahrnehmung der näheren Umgebung durch das Fahrzeug, um „natürliches“ vorausschauendes Fahren auch für Maschinen zu ermöglichen und um notwendige Basisdaten wie z. B. hochauflösende Straßenkarten effizient aktuell zu halten.

Fährt ein hochautomatisiertes Fahrzeug auf der Straße, nimmt eine Vielzahl von Kameras und weiteren Sensoren Informationen aus der Umgebung auf. Diese müssen in Echtzeit analysiert werden und ein zuverlässiges Ergebnis liefern, um beispielsweise Fußgänger auf der Fahrbahn zu erkennen. Bisher haben diese Aufgabe klassisch von Informatikern implementierte Algorithmen übernommen. Mit neuen Verfahren der Künstlichen Intelligenz insbesondere dem *Deep Learning*-Verfahren lernt der Computer anhand einer riesigen Menge von Daten zum Beispiel wie ein Mensch oder Baum aussieht. Durch die vielen Beispieldaten lernt er verschiedene Verkehrssituationen und kann diese nach Beendigung der Lernphase selbstständig im Fahrzeug anwenden. Der Trick dabei, die einzelnen Beispiele werden nicht auswendig gelernt, sondern die Algorithmen erkennen Muster und Gesetzmäßigkeiten in den Lerndaten und erkennen selbst dann Menschen, wenn diese im Winter dicke Mäntel tragen. ([→ Wie funktioniert maschinelles Lernen?](#))

Ähnliches passiert, wenn dem Computer das vorrauschaude Fahren – die Intuition – beigebracht wird. Hierbei sind

aber nicht Bilder, sondern Fahrsituationen die Beispieldaten für die Lernphase. Die große Herausforderung dabei: Wie können diese in großen Mengen gesammelt werden? Insbesondere wenn darin auch gefährliche Situationen mit enthalten sein sollen. Hier helfen synthetische Daten, die z. B. in Simulationen erzeugt werden. Mit Hilfe solcher künstlichen Trainingseinheiten kann dem Computer beigebracht werden, ohne menschliche Unterstützung Verkehrssituationen wahrzunehmen, sie zu verstehen und entsprechend darauf zu reagieren.

Trotz der Fortschritte durch die neuen Verfahren der Künstlichen Intelligenz ist der Mensch im Verkehr dem Computer immer noch voraus. Selbst in völlig neuen Umgebungen eines fremden Landes finden wir uns schnell zurecht. Der Computer braucht hier Unterstützung durch die hochgenauen Karten. Hier werden Basisdaten wie z. B. die genaue Position der Ampel-Signalgeber mit genutzt, damit das autonome Fahrzeug nicht das ganze Kamerabild absuchen muss, sondern gleich genau hinschauen kann. Diese hochgenauen Karten müssen dabei stets aktuell bleiben. Dabei helfen die eher statistischen Methoden der Künstlichen Intelligenz, etwaige Änderungen im Fahrzeug und über viele Fahrzeuge zusammengefasst sicher zu ermitteln – z. B. ob ein Verkehrsschild nur zugeschnitten ist oder abgebaut wurde.

**Dr. Ilja Radusch**  
Fraunhofer FOKUS

KI basiert auf Statistik. In der Wirtschaft wird KI eingesetzt, um Vorhersagen über Verhalten zu treffen, etwa darüber, ob sich jemand zum Kauf eines Produktes entschließen wird. KI wird auch genutzt, um zu bewerten, wie groß das Risiko eines unerwünschten Verhaltens ist, etwa, dass jemand einen Kredit nicht zurückzahlt. Dies beruht auf der Annahme, dass vergangenes auf künftiges Verhalten schließen lässt. Anhand von Verhaltensdaten werden Wahrscheinlichkeiten für die Zukunft berechnet.

Für die Bestimmung dieser Wahrscheinlichkeiten wird häufig ein bestimmtes KI-Verfahren verwendet: das maschinelle Lernen. Dabei wird ein Vorhersagemodell darauf „trainiert“, aus Verhaltensmustern und Korrelationen, die sich den Daten entnehmen lassen, möglichst treffsicher jene Muster herauszufiltern, die auf ein künftiges Verhalten hindeuten, das für die Betreiber von Interesse ist. (→ [Wie funktioniert maschinelles Lernen?](#)) So werden beispielsweise Menschen, die in einem Webshop mehrmals am Tag ein Produkt ansehen, dieses wahrscheinlich auch kaufen. Daten aus der Vergangenheit, etwa über Verhalten eines Kundenstamms, das tatsächlich mit getätigten Einkäufen in Verbindung stand, werden auf diese Weise verwendet, um Vorhersage-Modelle zu entwickeln.

Die Stärke der KI liegt darin, dass sie in sehr großen Datenmengen über verschiedene Verhaltensweisen relevante Muster erkennen kann. Dazu gehören beispielsweise Daten über die Verweildauer auf einer Produktseite, Mauszeigerbewegungen, die Anzahl der Aufrufe ähnlicher Produktseiten usw. (→ [Welche Bedeutung haben Daten für die KI?](#)) Zusätzlich werden die von der KI identifizierten Verhaltensmuster darauf getestet, in welchen Fällen die vorhergesagten Käufe tatsächlich eingetreten sind. So kann man das Modell immer weiter verfeinern und immer bessere Vorhersagen erhalten. Diese Verfeinerung kann zum Teil auch automatisiert werden, etwa indem die KI selbst überprüft, ob ein Kauf eingetreten ist, und entsprechend die Gewichtung der vorhersagerelevanten Verhaltensdaten anpasst. Daher spricht man von maschinellem Lernen. Die Vorhersagesoftware verbessert sich gewissermaßen von selbst.

Je vielfältiger die Trainingsdaten für so ein Verfahren sind, desto besser funktioniert es in der Praxis. Bei einem Training mit nicht-repräsentativen Daten über die Kundschaft eines Unternehmens, könnte die KI Verhaltensmuster identifizieren, die zwar in der Kundengruppe auftreten, von der die Trainingsdaten stammen, die aber bei einer breiteren Population gar keine Rolle spielen. Mit maschinellem Lernen kann eine KI also auch falsche Muster „erlernen“, die dann die Treffsicherheit ihrer Vorhersagen verringern.

**Philipp Otto / Eike Gräf**  
iRights.Lab

Noch vor wenigen Jahren war ein Navigationsassistent nur für Autos der Oberklasse verfügbar. Ein großer Bildschirm, eine CD oder DVD mit den notwendigen Kartendaten sowie die GPS-Verbindung mit dem Navigationssatelliten sorgten dafür, dass man sich den Blick auf die Straßenkarte sparen konnte. Alle paar Jahre musste ein neuer Datenträger mit den aktuellsten Straßenkarten gekauft werden; Daten über die Verkehrslage wurden über das angeschlossene Autoradio übertragen.

Nicht nur, dass nun jeder Smartphone-Nutzer ein Navigationsgerät mit sich führt; er liefert auch gleich die notwendigen Daten, die diese Software benötigt, um Verkehrsströme zu messen und gegebenenfalls auch lenkend einzugreifen. Jeder, der ein solches Gerät einsetzt, sorgt dafür, dass u.a. die genaue Position und Geschwindigkeit übermittelt werden. Die Software erkennt so in Echtzeit Staus oder Sperrungen und kann alternative Routen vorschlagen. Liefert der Nutzer noch per Eingabe den Grund für die Verzögerung (z. B. Unfall oder Grenzkontrollen), erfahren sofort alle anderen Verkehrsteilnehmer, was sich auf der Straße vor ihnen tut.

Dabei ist die Wegführung als Reaktion auf Befehle des Fahrers nur der Anfang der intelligenten Navigation. Autohersteller arbeiten an Systemen, die vorausschauend aktiv werden und somit das Auto auf zukünftige Ereignisse konditionieren. Dazu muss das Navigationssystem zunächst

lernen. Jeder Fahrer erhält ein Profil, das Informationen zu dessen Fahrten speichert. Ziele, Abkürzungen und Schleichwege, aber auch Uhrzeit und beispielsweise die Sitzbelegung können hier als Informationen dienen. So ist an Wochentagen zu bestimmten Uhrzeiten die Fahrt zur Arbeit am wahrscheinlichsten. Sitzt auf dem Rücksitz ein Kind, plant die Navigation automatisch den besten Weg zu Kindergarten oder Schule ein. Spannend wird es dann, wenn das Navigationssystem mit den fahrzeuginternen Systemen vernetzt wird und somit z. B. energiesparend auf die Fahrweise Einfluss nehmen kann.

Ziel ist es, dass sich Autos ganz ohne das Eingreifen eines Fahrers über die Straßen bewegen. Mit Hilfe von Videokameras, Ultraschall, Radar- und Laserscannern sowie jeder Menge Sensoren, ist das sogenannte autonome Fahrzeug schon heute teilweise unterwegs. Diese Technologien erfassen dauerhaft das Geschehen und tasten alle paar Millisekunden die Umgebung rund um das Auto ab. Im bord-eigenen Computer werden alle Informationen der einzelnen Bausteine zu einem Gesamtbild verarbeitet und mit unzähligen dort gespeicherten Szenarien verglichen. Das System reagiert entsprechend dieses Gesamtbildes auf die jeweilige Situation. Die Vernetzung aller Fahrzeuge sowie der Datenaustausch untereinander ermöglichen dann ein nahezu perfektes Bild der Verkehrssituation.

**Stefan Gehrke**  
bfn

Wie soll ich handeln? Gibt es Normen, an denen ich mich orientieren kann? Als Aristoteles den Begriff der Ethik eingeführt hatte, bezog er all diese Fragen ausschließlich auf den Menschen. In den letzten Jahrzehnten haben sich Maschinen allerdings stark verändert: Ein bestimmtes Verhalten kann programmiert werden, sie können auf Basis riesiger Datenmengen lernen und in definierten Situationen auch selbstständig entscheiden. Reicht das schon aus, um moralisch zu handeln? Und wessen Aufgabe ist es nun festzulegen, wann Maschinen selbstständig entscheiden dürfen?

Menschen müssen festlegen, wie Maschinen agieren sollen. Informatiker werden dabei versuchen, ethische Regeln für die Maschinen zu übersetzen, so diese denn mit der aktuellen Technologie umsetzbar sind. Derartige ethische Regeln zu definieren, sollte in einem gesellschaftlichen Diskurs stattfinden. Einen ersten Schritt in diese Richtung machte beispielsweise Ende 2016 eine aus Juristen, Philosophen, Theologen, Ingenieuren, Datenspezialisten und Verbraucherschützern bestehende Ethikkommission, die ethische Leitlinien für die Programmierung automatisierter Fahrsysteme entwickeln sollte. Im Juni 2017 wurde ihr Bericht vorgestellt. Das Fazit: Automatisierte Autos können für mehr Sicherheit sorgen, in Grenzsituationen, in Dilemmata, in denen es ausweglos um Leben oder Tod geht, sind sie aber überfordert. In solchen Momenten sind Entscheidungen von der konkreten tatsächlichen Situation unter Einschluss unberechenbarer Verhaltensweisen Betroffener abhängig. Sie sind deshalb nicht eindeutig normierbar und auch nicht ethisch zweifelsfrei programmierbar, zumal wenn gar nicht die Handlung selbst programmiert wird, sondern maschinelles Lernen ([→ Wie funktioniert maschinelles Lernen](#)) eingesetzt wird. Als zentrale Leitlinien für die Programmierung definierten sie, dass der Schutz menschlichen Lebens immer Vorrang hat, etwa gegenüber Sach- oder Tierschäden, und jedes Menschenleben gleichwertig ist. Bei unausweichlichen Unfallsituationen dürfe es also keine Qualifizierung von Menschen nach persönlichen Merkmalen wie Alter, Geschlecht, körperlicher oder geistiger Konstitution geben.

Autonome Autos können aber nur dann für einen sicheren Straßenverkehr sorgen, wenn die Bürgerinnen und Bürger die Grenzen des automatisierten Autos kennen und berücksichtigen. Beispielsweise wird ein Fußgänger, der unbedacht und in Erwartung der „rücksichtsvollen“ autonomen Autos auf die Straße tritt, weiterhin den Verkehr gefährden. Es ist eine Art „Netiquette“ für den Umgang mit diesen Maschinen notwendig.

Ähnlich wie beim Datenschutz oder dem Hippokratischen Eid in der Medizin sind umsetzbare ethische Leitlinien für das Programmieren intelligenter Maschinen unerlässlich. Sie fördern das Vertrauen in die Technologie und damit auch den praktischen Einsatz der Supermaschinen.

**Dr. Ilja Radusch**  
Fraunhofer FOKUS

KI wird unter anderem eingesetzt, um die Wahrscheinlichkeiten zukünftiger Ereignisse einzuschätzen, um Zusammenhänge und Muster in großen Datensätzen zu erkennen sowie ganz allgemein um große Mengen an Informationen zu sortieren und zu kategorisieren. All diese Anwendungsarten helfen Menschen dabei, Situationen zu verstehen, zu bewerten und letztendlich Entscheidungen zu treffen.

In vielen Kontexten wird KI ganz gezielt als Entscheidungshilfe eingesetzt, beispielsweise in der medizinischen Diagnostik oder bei der Vergabe von Krediten. In den USA gibt es zudem ein viel diskutiertes Anwendungsbeispiel, das darin besteht, Entscheidungen über die Freilassung Inhaftierter zu treffen. Ob diese auf Bewährung freikommen, wird mit Hilfe eines Risiko-Wertes entschieden, der durch eine KI berechnet wird. In Zusammenhang mit Entscheidungen, die für betroffene Personen besonders bedeutsam sind, fordern viele Experten, dass diese von Menschen getroffen werden. In Europa ist sogar gesetzlich geregelt, dass rechtlich bindende oder vergleichbar folgenreiche Entscheidungen über Menschen nicht vollautomatisch getroffen werden dürfen.

Doch auch wenn eine KI lediglich eine Empfehlung generiert oder einen Wahrscheinlichkeitswert bezüglich eines bestimmten Risikos ausgibt, ist der Einfluss der Technologie auf die jeweilige Entscheidung potenziell sehr hoch. Das hat verschiedene Gründe.

Einerseits werden Computerprogramme häufig als objektiv wahrgenommen. Ihre Ausgaben treffen daher auf eine hohe Akzeptanz, obwohl auch Computerprogramme durchaus „ungerechte“ Ergebnisse produzieren können, zum Beispiel wenn sie Daten verarbeiten, die gesellschaftliche Vorurteile widerspiegeln.

Weiterhin braucht es ein hohes Selbstbewusstsein bei jenen Menschen, die sich gegen die Empfehlung einer KI entscheiden und diese Entscheidung im Zweifel gegenüber Ihren Arbeitgebern rechtfertigen müssen. Dies gilt ganz besonders, wenn so eine Entscheidung entgegen einer KI-generierten Empfehlung unerwünschte Folgen hat, wie etwa die Freilassung einer Person, die wenig später wieder straffällig wird.

Letztlich besteht auch dann eine starke Beeinflussung von Entscheidungen durch KI, wenn die Ausgabe einer KI an die Stelle von Argumenten und Informationen tritt, die bis dato als Entscheidungsgrundlage gedient haben. Eine KI mag Zusammenhänge in komplexen und umfangreichen Datensätzen erkennen und bewältigen können, Menschen können dies häufig jedoch nur eingeschränkt nachvollziehen. Hierin besteht aktuell eine Herausforderung für den legitimen Einsatz von KI in Entscheidungsprozessen.

**Philipp Otto / Eike Gräf**  
iRights.Lab



Künstliche Intelligenz stellt die Denkweise des Menschen auf den Kopf. Menschen bewältigen Lebenssituationen, indem sie sich ein Bild der Lage machen, in Gedanken eine Hypothese entwickeln, was zu tun ist, und dann im Handeln ausprobieren, ob ihre Überlegung ein praktikabler Weg zur Zielerreichung ist.

Dieser meist unbewusst ablaufende Vorgang wird in den empirischen Wissenschaften zur wichtigsten Methode: Hypothesen werden entwickelt, die im anschließenden Experiment geprüft werden. Widersprechen ihnen die Versuchsergebnisse, müssen sie verworfen werden. Ansonsten gelten sie als richtig. In den empirischen Wissenschaften spielt Kausalität, der ursächliche Zusammenhang zwischen zwei Ereignissen, eine zentrale Rolle: Sie trägt maßgeblich zur Aussagekraft einer wissenschaftlichen Theorie bei. Korrelationen, das gleichzeitige Auftreten zweier Ereignisse, allein reichen dagegen nicht. Sie können zufällig sein und zu falschen Schlussfolgerungen führen.

Durch KI verändert sich Wissenschaft. Ein neuer Weg, Forschungsergebnisse zu generieren, kommt hinzu. Die herausragende Leistungsfähigkeit der KI macht es möglich, in großen Datenmengen Muster zu finden, die mit herkömmlichen Methoden nicht erkennbar sind, und daraus Schlussfolgerungen zu ziehen. Ein Beispiel ist die Genomforschung. Die Genome vieler Lebewesen sind sequenziert. Nun gilt es, in der großen Fülle genetischer Daten (alleine das menschliche Genom besteht aus 3,2 Milliarden Bausteinen!), Strukturen zu erfassen, die für die Funktionsweise der Lebewesen von Bedeutung sind. KI hilft, Gene und ihre Regulationseinheiten zu identifizieren, um daraus Erkenntnisse, z. B. für die Medizin, abzuleiten. Revolutionär ist, dass nicht mehr unbedingt eine Hypothese am Anfang der Forschung steht, sondern empirische Daten: Das Stochern im (Daten-) Heuhaufen lohnt sich, KI findet die Nadel!

Auch im gesellschaftlichen Bereich hat dieses Vorgehen gravierende Folgen: KI, die z. B. darauf trainiert ist, Kriminalität zu erkennen, könnte potentiell Kriminelle identifizieren – bevor sie eine Straftat begangen haben. Dies hat nicht nur positive Auswirkungen: Falsche Verdächtigungen und Diskriminierungen könnten die Folge sein – weil KI wohl Korrelationen identifiziert, aber keinen Nachweis für Ursache-Wirkungs-Zusammenhänge liefert. Deshalb ist es wichtig, die Software der KI und die Daten, mit denen KI trainiert wird, sorgfältig zu prüfen und sich nicht blindlings auf sie zu verlassen.

**Dr. Norbert Arnold**  
Konrad-Adenauer-Stiftung

### Lesetipp:

**Wie Maschinen denken lernen:** <https://www.zeit.de/2018/12/viktor-mayer-schoenberger-digitalisierung-big-data> (aufgerufen am 16.08.2018)

KI-Programme existieren in der Theorie seit langer Zeit. Viele Verfahren wurden bereits vor Jahrzehnten entwickelt. Ein Hauptgrund dafür, dass aktuell beim Einsatz von KI große Fortschritte gemacht werden, besteht darin, dass erst heute die benötigten Datenmengen (Big Data) zur Verfügung stehen und die entsprechende Rechenleistung existiert, um sie zu verarbeiten.

Die meisten KI-Anwendungen funktionieren auf der Grundlage maschinellen Lernens. (→ *Wie funktioniert maschinelles Lernen?*) Algorithmen werden mit für die jeweilige Anwendung relevanten Daten „trainiert“, um in diesen Daten enthaltene Muster automatisiert und sicher zu erkennen. So funktioniert beispielsweise eine KI zur Bilderkennung umso verlässlicher, je mehr Bilder ihr als Trainingsdaten zugeführt werden. (→ *Wie funktioniert Bilderkennung mit KI?*) Eine KI, die eine Prognose über die Kaufbereitschaft von Kunden ermitteln soll, wird umso treffsicherer, je mehr Verhaltensdaten von bisherigen Kunden eingegeben werden, in denen sie nach relevanten Wechselbeziehungen suchen kann.

Neben der verfügbaren Menge hat die Qualität (Aktualität, Genauigkeit, Richtigkeit, etc.) und die Zusammensetzung der genutzten Daten eine Bedeutung für das Funktionieren von KI-Anwendungen. Veraltete, ungenaue, falsche oder unpassende Daten beeinflussen die Wirkungsweise einer KI und damit die Qualität der Ergebnisse. Die Verwendung

von Datensätzen, die das Anwendungsgebiet unvollständig, verzerrt oder selektiv abbilden, führen zu einer Verzerrung (Bias) beim maschinellen Lernen und der Qualität der Ergebnisse. Wenn etwa eine KI, die das Auftreten von Krankheitssymptomen analysiert, nur mit den Daten einer bestimmten sozialen Gruppe trainiert wird, ist nicht auszuschließen, dass die Ergebnisse in Bezug auf andere soziale Gruppen und ihre abweichenden Lebensbedingungen wenig brauchbar sind. (→ *Wie wird KI bei Vorhersagen und Bewertungen eingesetzt?*)

Nicht nur der eingesetzte Algorithmus, sondern auch die Daten haben somit einen Einfluss auf Wirkungsweise und die Ergebnisse einer KI. Ein Verständnis für die Daten, mit denen eine KI entwickelt wurde oder arbeitet, ist also von hoher Bedeutung. Daher gilt es, technische oder konzeptionelle Mängel bei der Verarbeitung sowie kulturelle, politische oder soziale Begrenzungen der Aussagekraft der verwendeten Daten offenzulegen, zu reflektieren und zu bewerten. Nur so können Verzerrungen, Diskriminierungen und Fehlinterpretationen durch KI eingeschätzt werden. Gerade der Einsatz von KI in sozialen Zusammenhängen, beispielsweise zur Vorhersage und Bewertung (Kreditwürdigkeit, Scorings, Entwicklungsprognosen, etc.) unterliegt daher einer hohen ethischen und politischen Verantwortung.

**Philipp Otto / Eike Gräf**  
iRights.Lab

Ein großer Vorteil von KI-Anwendungen besteht darin, in großen Datenmengen komplexe Muster und Zusammenhänge entdecken zu können. Sie können Informationsmengen verarbeiten, die für Menschen nicht mehr zu bewältigen sind. Je nach Art der Anwendungen können die Ergebnisse einer KI dadurch allerdings auch schwer nachzuvollziehen sein.

Im Bereich der Bilderkennung werden etwa KI-Verfahren eingesetzt, die auf maschinellem Lernen beruhen. (→ *Wie funktioniert maschinelles Lernen?*) Eine KI, die lernen soll, Hunde von Katzen zu unterscheiden, bekommt in der Regel tausende von Bildern als Trainingsdatensatz eingespeist. Dabei sind die Motive eindeutig als Hunde oder Katzen gekennzeichnet. Auf dieser Basis sucht die KI nun selbst einen Weg, Bildelemente zu identifizieren, die eher auf einen Hund oder eine Katze hindeuten. Sie „kennt“ sozusagen die richtige Antwort, aber nicht den Weg, diese zu erlangen. Mit Hilfe von Bildern, die nicht im Voraus von Menschen kategorisiert wurden, lässt sich dann prüfen, ob die KI in der Lage ist, die Bilder richtig zuzuordnen. Selbst wenn die KI dies zuverlässig tut, ist jedoch nicht immer klar, worauf diese Zuordnung beruht. Während Menschen einen Hund an seinem typischen Aussehen erkennen, legt eine KI ihrer Zuordnung möglicherweise das Auftreten bestimmter Farbkontraste, oder Bildkanten zugrunde, die Menschen gar nicht wahrnehmen.

Um dieser Herausforderung bei der Nachvollziehbarkeit zu begegnen, wird an erklärbarer KI geforscht. Ein Weg, die Funktionsweise von Bilderkennungsverfahren transparenter zu machen, besteht beispielsweise darin, Bildelemente hervorzuheben, die zu einer bestimmten Klassifizierung geführt haben. Wenn eine KI zur „Erklärung“ bestimmte Bildbereiche einfärbt und dabei deutlich wird, dass vor allem Bilder, die eine Leine enthalten, zur Klassifizierung „Hund“ geführt haben, dann lässt sich nachvollziehen, dass die KI zwar in vielen Fällen korrekte Ergebnisse produzieren wird, aber nicht aus Gründen, die Menschen für belastbar und zuverlässig halten.

Während dies bei der Klassifizierung von Hunden und Katzen relativ harmlos anmutet, ist es in vielen Anwendungsfällen von entscheidender Bedeutung, die Gründe, aus denen eine KI zu einem bestimmten Ergebnis kommt nachvollziehen und kritisch hinterfragen zu können. Im Bereich der Bilderkennung ist zum Beispiel die Radiologie ein solcher Anwendungsfall. Die Entwicklung transparenter und nachvollziehbarer KI-Verfahren ist auch über die hier als Beispiel verwendete Bilderkennung in vielen Bereichen eine Voraussetzung für den sicheren Einsatz von KI. (→ *Kann man Ethik programmieren?*)

**Philipp Otto/Eike Gräf**  
iRights.Lab

KI hat wie jede andere Technologie einen ambivalenten Charakter. Sie bietet Vorteile, die dem Menschen von Nutzen sind, birgt aber auch Risiken in sich, die minimiert werden müssen. Ein „Nullrisiko“ gibt es nicht. Entscheidend ist, Risiken und Nutzen in einem verantwortbaren Verhältnis zu halten. Welche Risiken akzeptabel sind und in Kauf genommen werden können, um vom Nutzen neuer Technologien zu profitieren, muss die Gesellschaft entscheiden. Die Akzeptabilität muss immer wieder neu, öffentlich und transparent verhandelt werden. Dies ist eine zentrale Aufgabe von Gesellschaft und Politik.

Die Ambivalenz der KI ist offensichtlich: Sie ist ein gewaltiger Innovationsmotor, verändert aber zugleich die Erwerbsarbeit. Sie führt zu höheren Unternehmensgewinnen, kann aber auch zur Ungleichheit beitragen. Sie eröffnet neue Informations- und Kommunikationswege, denen Menschen, die mit diesen Techniken nicht umgehen, aber nicht gewachsen sind. Viele weitere Beispiele ließen sich ergänzen. Es kommt also darauf an, Chancen und Risiken der KI nicht schicksalhaft hinzunehmen, sondern gesellschaftlich und politisch zu gestalten.

Jeder steht vor der Aufgabe, seine digitalen Kompetenzen zu verbessern – nicht nur in der Bedienung neuer intelligenter Technik: smarte Assistenten zu Hause, Autopiloten in Fahrzeugen, Suchmaschinen bei Internetrecherchen usw.

Es braucht zudem kulturelle Kompetenzen, um KI sinnvoll zu nutzen und sie mit individuellen Bedürfnissen und Interessen in Einklang zu bringen. Und schließlich wird von jedem kritische Urteilsfähigkeit erwartet, um im Umgang mit KI entstehende Gefahren und Diskriminierungen zu erkennen.

Politik sollte die Menschen bei den durch KI hervorgerufenen Veränderungen nicht allein lassen. Es gilt, die richtige Balance zwischen Technologieförderung und Technologieregulierung zu finden. Dabei ist wissenschaftlicher Rat von Nutzen. Das Bildungssystem diesen Gegebenheiten anzupassen, ist wichtiger denn je. Digitalisierung und KI müssen Bildungsthema im gesamten Lebensverlauf werden.

KI erlebt derzeit einen medialen Hype und dominiert öffentliche Debatten. Erstrebenswert ist eine Normalisierung des Umgangs mit ihr, die sie weder als Teufelswerk, noch als Allheilmittel erscheinen lässt.

**Dr. Norbert Arnold**  
Konrad-Adenauer-Stiftung

#### Lesetipps:

Manuela Lenzen: **Künstliche Intelligenz: Was sie kann & was uns erwartet.** Verlag C. H. Beck, München 2018.

Max Tegmark: **Leben 3.0. Mensch sein im Zeitalter von Künstlicher Intelligenz.** Ullstein-Verlag, Berlin 2017.