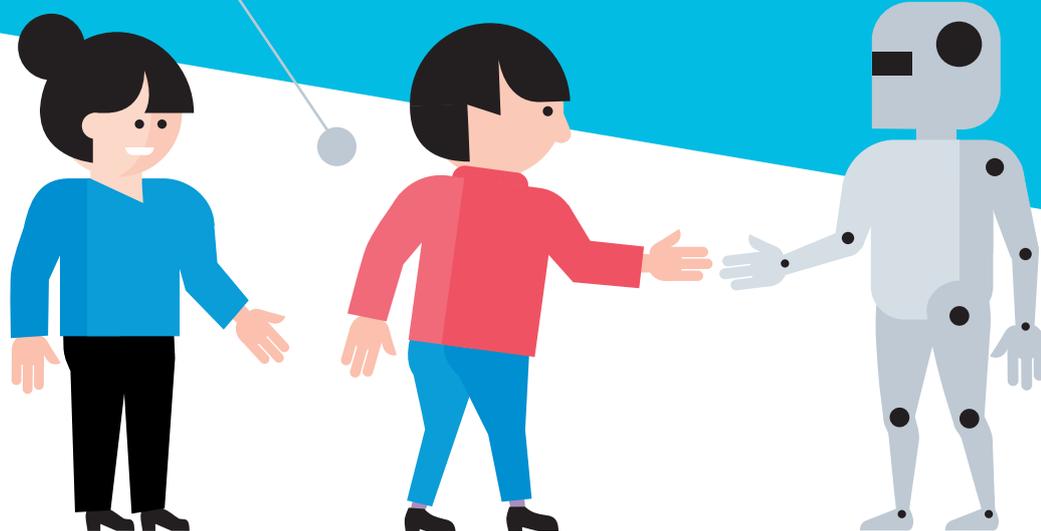


Kopflös klug

Kurzfassung der Studie «Wenn Algorithmen für uns entscheiden:
Chancen und Risiken der künstlichen Intelligenz» von TA-SWISS



Die Stiftung TA-SWISS, ein Kompetenzzentrum der Akademien der Wissenschaften Schweiz, setzt sich mit den Chancen und Risiken neuer Technologien auseinander.

Die hier vorliegende Kurzfassung basiert auf einer wissenschaftlichen Studie, die im Auftrag von TA-SWISS von einem interdisziplinären Projektteam unter der Leitung von Markus Christen (Digital Society Initiative Universität Zürich), Clemens Mader (Eidgenössische Materialprüfungs- und Forschungsanstalt Empa) und Johann Čas (Österreichische Akademie der Wissenschaften ÖAW) durchgeführt wurde. Die Kurzfassung stellt die wichtigsten Resultate und Schlussfolgerungen in verdichteter Form dar und richtet sich an ein breites Publikum.

Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz

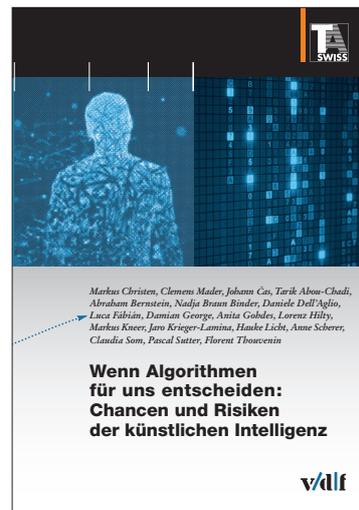
Markus Christen, Clemens Mader, Johann Čas, Tarik Abou-Chadi, Abraham Bernstein, Nadja Braun Binder, Daniele Dell'Aglio, Luca Fábíán, Damian George, Anita Gohdes, Lorenz Hilty, Markus Kneer, Jaro Krieger-Lamina, Hauke Licht, Anne Scherer, Claudia Som, Pascal Sutter, Florent Thouvenin

TA-SWISS, Stiftung für Technologiefolgen-Abschätzung (Hrsg.).

vdf Hochschulverlag an der ETH Zürich, 2020.
ISBN 978-3-7281-4001-2

Die Studie steht als eBook zum freien Download bereit: www.vdf.ch

Die vorliegende Kurzfassung ist ebenfalls online verfügbar: www.ta-swiss.ch



Künstliche Intelligenz in aller Kürze	4
Ihre Chancen ...	4
... ihre Risiken ...	4
... und einige Empfehlungen	5
Künstliche Intelligenz, ein Kind der Digitalisierung	5
Vom mathematischen Problem zum komplexen Modell	5
Den Menschen überflügeln	6
Verzerrungen, mangelnde Fairness und Vertrauensdefizite	7
Vielfältige Anwendungsfelder	8
KI und Konsum: Individualisiert auf Einkaufstour	8
Persönliche Vorlieben als Orientierungsmarken	8
Mensch oder Maschine?	8
Digitale Leimruten	9
Schutz der Daten schützt auch die Person	9
Schwierige Balance zwischen Personalisierung und Datenschutz	10
KI und öffentliche Kommunikation: Meinungsmache im medialen Echoraum	10
Bestärkt in der Blase	10
Verschörungstheorien aus der Trollfabrik	11
Demokratie in Gefahr	11
Skeptischer Blick auf soziale Netzwerke	11
KI in der Berufswelt: Wenn Maschinen Kopfarbeit verrichten	12
Karriere unter der Schirmherrschaft von KI	12
Arbeit im maschinellen Takt	13
Massnahmen gegen die Zweiteilung des Arbeitsmarktes sind gefragt	13
KI in Bildung und Forschung: Unterstützung für die Humanintelligenz	14
Jedem Kind seine Superlehrkraft	14
Auf der Lauer nach den Spitzenkräften der Zukunft	14
Wachsender Einfluss von Privatunternehmen auf die Bildung	15
KI wird erforscht – und bringt Forschung weiter	15
Schlüsselkompetenzen ausbauen und Transparenz sicherstellen	15
KI in der Verwaltung: Hoheitliche Befugnisse für künstliche Intelligenz?	16
Im Einsatz gegen Betrug und Verbrechen	16
Grundrechte weisen staatliche Willkür in die Schranken	16
Zurückhaltende Schweiz	17
Keine Delegation von Entscheidungsgewalt an die Maschine	17
«Kluge» Maschinen – eine Herausforderung für geltendes Recht	18
KI-System an die Leine legen?	18
Kreativität aus erster und zweiter Hand	18
Fairness lässt sich schwer berechnen	19
Ethikrichtlinien auf europäischer Ebene	19
Verzerrungen und Intransparenz fallen ins Gewicht	19
Empfehlungen: Ein weites Feld gestalten	20
Spezifische Anwendungen ausleuchten	20
Hohe Anforderungen an die Behörden	20
Transparenz und verdaubare Information	20
Gütesiegel statt generelle Marktzulassungen	20
Nicht bloss auf technische Fragen fokussieren	20

Künstliche Intelligenz in aller Kürze

Im Unterschied zum herkömmlichen Computerprogramm, das genau festgeschriebene Aufgaben nach vorgegebenen Schritten löst, wirkt künstliche Intelligenz (KI) kreativ. Viele KI-Systeme sind in der Lage, dank eines Trainings an grossen Datensätzen, versteckte Regelmässigkeiten zu erkennen und Lösungen vorzuschlagen, auf die die Anwender nicht von selbst gekommen wären. Fortgeschrittene KI-Systeme vereinigen verschiedene Fähigkeiten wie die Verarbeitung von Sprache, das Ordnen von Information sowie maschinelles Lernen und sind in der Lage, sogar anspruchsvolle Tätigkeiten zu verrichten, für die zurzeit noch gut ausgebildete Berufsleute zuständig sind.

Ihre Chancen ...

KI-Systeme lassen sich in den unterschiedlichsten Anwendungsfeldern einsetzen. Sie erledigen vieles schneller und oft auch präziser als Menschen und könnten diese bei komplexen Aufgaben in die Lage versetzen, verschiedene Tätigkeiten wesentlich effizienter als bisher auszuführen.

Oft ermöglichen es KI-Systeme, Dienstleistungen und Angebote besser an die Bedürfnisse und Fähigkeiten der einzelnen Individuen anzupassen. In der Personalisierung der Angebote liegt eine grosse Stärke von KI.

... ihre Risiken ...

Viele KI-Systeme müssen mit riesigen Datensätzen trainiert werden, damit sie die von ihnen gewünschten Fertigkeiten erlangen. Privatsphäre und Datenschutz laufen Gefahr, vom immensen Datenbedarf der KI ausgezehrt zu werden.

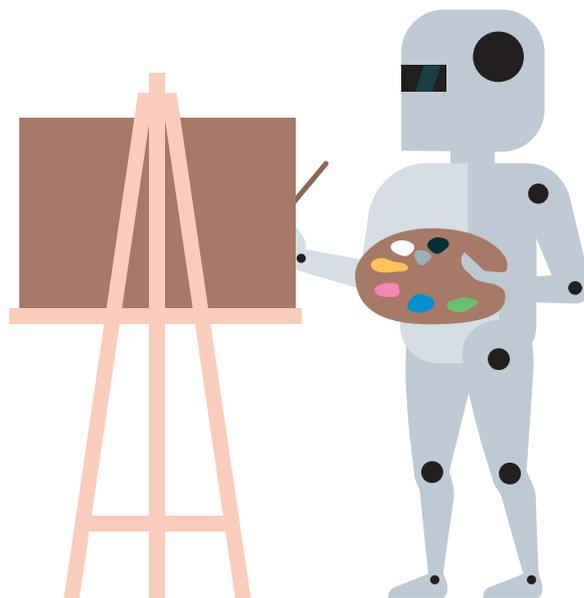
Wenn die der KI zugrundeliegenden Datensätze Fehler enthalten, sind auch die Ergebnisse falsch. Problematisch sind zudem unausgewogene Daten, die zwar zu mathematisch korrekten, inhaltlich aber verzerrten Resultaten führen, die unter Umständen bestimmte Personengruppen systematisch benachteiligen.

Selbstlernende KI-Systeme entwickeln sich weiter und kommen unter Umständen zu Ergebnissen, die für ihre Urheber, vor allem aber auch für die Betroffenen nicht mehr nachvollziehbar sind.

... und einige Empfehlungen

Der Einsatz von KI-Systemen sollte deutlich und auf einfache Weise kenntlich gemacht werden, damit Betroffene wissen, dass die Interaktion mit einem KI-System stattfindet und nicht mit einem Menschen.

Wichtige Entscheidungen, die Personen betreffen, sollen nicht ohne reifliche Abwägungen der Vor- und Nachteile einem KI-System überlassen werden. Geht



es um relevante personenbezogene Fragen, muss in der Regel das Ergebnis des Systems durch einen Menschen nachgeprüft und begründet werden.

Der Einsatz von KI hat Folgen, die weit über die technischen Implikationen hinausreichen. Wer KI entwickelt, einsetzt oder mit ihren Ergebnissen weiterarbeitet, sollte daher auch über Kenntnisse in Ethik und Recht verfügen und darüber hinaus zu interdisziplinärer Zusammenarbeit mit Vertretern anderer Wissenschaftsgebiete bereit und fähig sein.

Methodisch beruht die Studie in einem ersten Schritt auf einer breit angelegten Literaturanalyse. Diese lieferte die Grundlage, um in einer zweiten Phase Fachleute in mehreren Durchgängen nach ihrer Einschätzung der herausgearbeiteten Problemfelder und Thesen zu befragen. Die Ergebnisse aus Literaturanalyse und Umfragen mündeten in Empfehlungen für verschiedene Adressaten.

Künstliche Intelligenz, ein Kind der Digitalisierung

Künstliche Intelligenz (KI) ist längst keine Science Fiction mehr, sondern schleichend Teil unseres Alltags geworden. Grund genug, um nicht nur ihren Nutzen wie die gesteigerte Effizienz zu würdigen, sondern auch ihre Schattenseiten – insbesondere ihren Hunger nach Daten – auszuleuchten.

Geht man von der Herkunft des Wortes «Intelligenz» aus, steht diese für die geistige Freiheit von Lebewesen, die keinen vorgegebenen Pfaden folgen, sondern sich Alternativen ausdenken und sich für diese entscheiden können. Doch wofür genau steht KI, die Abkürzung der Wortkombination «künstliche Intelligenz»?

KI begann 1950 in den Fokus der Wissenschaft zu rücken. Damals hielt der britische Computerwissenschaftler Alan Turing in einem Artikel der Zeitschrift «Mind» fest, von einer intelligenten Maschine könne dann die Rede sein, wenn diese eine Unterhaltung mit einem Menschen so perfekt zu imitieren vermöge, dass er sich im Dialog mit einer realen Person wähne. Der sogenannte Turing-Test – er wird am Bildschirm mit einer Tastatur und ohne Sicht- oder Hörkontakt zum Gesprächspartner durchgeführt – gilt seit der ersten wissenschaftlichen Konferenz über KI, die 1956 in Dartmouth (USA) stattfand, als Lackmustest für die Intelligenz von künstlichen Systemen. An dieser Tagung wurde auch der Ausdruck «künstliche Intelligenz» geprägt.

Vom mathematischen Problem zum komplexen Modell

Seit ihren Anfängen ist die KI als Forschungsgegenstand der Computerwissenschaft erheblich vielschichtiger geworden. Während zunächst Lösungsalgorithmen für mathematische und geometrische Probleme im Vordergrund standen, rückte eine Annäherung an komplexere Prozesse erst ab den 1990er-Jahren dank der zunehmenden Rechnerleistung der Computer in Griffnähe. Um nämlich bedeutungsvolle Muster aus akustischen Signalen, Bildern oder anderen Eingaben herausarbeiten zu können, müssen viele KI-Systeme mit einer riesigen Menge an Daten trainiert werden. Die um sich greifende Digitalisierung und damit die wachsende Verfügbarkeit grosser Datenvolumen liefern eine wichtige Grundlage für KI.

Heute umfasst KI unterschiedliche wissenschaftliche Untersuchungsgebiete, in denen versucht wird, mithilfe eines Artefakts – also etwa eines Computerprogramms – rationales oder menschliches Handeln nachzubilden. So zielt eine Untergruppe von Arbeiten darauf ab, das Denken und Lernen von Menschen zu reproduzieren. Nebst denjenigen Verfahren, die sich grosser Datenmengen bedienen, um mithilfe von Algorithmen Regelmässigkeiten zu ermitteln und Modelle zu entwerfen, kommen auch Verfahren zum Einsatz, die sich auf Expertenwissen

oder logische Denkschritte abstützen. Fortgeschrittene KI-Systeme vereinigen verschiedene Fähigkeiten wie die Verarbeitung von Sprache, das Ordnen von Information und maschinelles Lernen.

In den letzten Jahren hat KI viel dazu beigetragen, Komfort und Sicherheit in unserem Alltag zu steigern: Mobilitätskonzepte, die den öffentlichen Verkehr mit dem Mietauto oder dem Privatwagen verbinden, beruhen auf KI-Systemen, Eisenbahngesellschaften erhöhen ihre Sicherheit durch intelligente Schadmustererkennung, und die industrielle Fertigung nutzt KI für effizientere Abläufe – um nur einige wenige Beispiele zu nennen. Grundlagenforschung zur Beschaffenheit des Kosmos, wie sie das CERN betreibt, wäre ohne KI-gestützte Auswertung der riesigen Datenmengen unmöglich, und auch Simulationen des Klimawandels bedienen sich der KI. Verwenden KI-Systeme solche sachbezogenen Daten, ist das in aller Regel unproblematisch und erbringt oft erheblichen gesellschaftlichen Nutzen. Heikel wird es, wenn die KI mit personenbezogenen Daten arbeitet. Solche Anwendungen stehen denn auch im Blickpunkt der Studie von TA-SWISS.

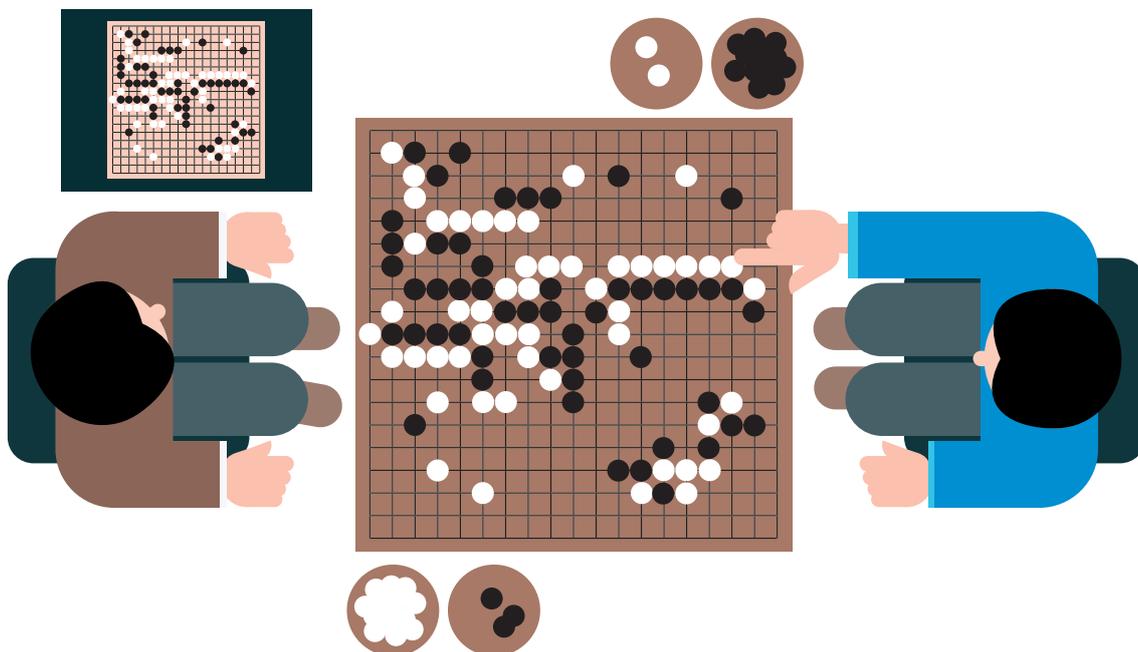
Den Menschen überflügeln

Insbesondere KI-Systeme, die auf maschinellem Lernen und auf neuronalen Netzen beruhen, lösen Probleme, an denen herkömmliche Computerprogramme scheitern. Bei anspruchsvollen und mehrdeutigen Aufgaben wie Übersetzungen, Bilderken-

nung und -analyse sowie Spielen vermag KI heute mit dem Menschen nicht nur mithalten, sondern ihn zuweilen gar zu übertreffen.

Ein Meilenstein wurde im März 2016 gesetzt, als eine KI im komplexen asiatischen Strategiespiel «Go» den damals weltbesten Spieler besiegte. Damit nicht genug: In der Folge arbeitete die Softwarefirma «DeepMind» – eine Google-Tochter – weiter an ihrem Programm und entwickelte eine Version, die das Spiel allein aufgrund der Spielregeln selber lernte, statt wie die früheren Versionen von der Analyse unzähliger Stellungen aus Partien mit starken menschlichen Spielern auszugehen. Beginnend mit zufällig gespielten Partien auf Anfängerniveau, erreichte die «Alpha Go Zero» genannte Software innerhalb von drei Tagen Profi-Spielstärke und übertraf seine Vorgängerin, die den Weltranglisten-Ersten geschlagen hatte.

Die sich rasant entwickelnden KI-Systeme sind zunehmend in der Lage, Tätigkeiten auszuführen, die zuvor Menschen vorbehalten waren und werden damit zur Basistechnologie für unzählige Anwendungen. Kommen neue Formen der KI zum Einsatz, verändert sich die Rolle, die der Mensch bei der Definition und Behebung von Problemen spielt: Bislang war er es, der Fragestellungen und Aufgaben erkannte, entsprechende Handlungsvorschriften zur Lösung des jeweiligen Problems entwarf (und allenfalls Korrekturen nach ersten Fehlschlägen vornahm) und die Anweisungen schliesslich in relativ genau umschriebenen Anwendungsfeldern umsetzte. Ergab



die Evaluation, dass die vollzogenen Lösungsschritte nicht zum Erfolg geführt hatten, nahm wiederum der Mensch die erforderlichen Korrekturen vor. Im Unterschied dazu geht maschinelles Lernen weitgehend ohne menschliches Zutun vonstatten: Grosse Datenbestände strukturieren die Problematik vor, die Handlungsvorschriften zur Problemlösung – der sogenannte Algorithmus – werden aufgrund der Muster in den Daten und durch verstärkendes Lernen der Maschine entwickelt. Auch die Rückkopplung nimmt die Maschine eigenständig vor, indem der Algorithmus gegebenenfalls aufgrund weiterer, autonom erfasster Daten angepasst und präzisiert wird. Der Mensch bleibt in weiten Teilen des Prozesses aussen vor, auch wenn es ihm vorbehalten bleibt, das Ergebnis des KI-Systems gegebenenfalls zu verwerfen.

Was reichlich abstrakt klingt, zieht handfeste Folgen im täglichen Leben nach sich – etwa bei der Kreditvergabe von Banken: Im herkömmlichen Verfahren prüft eine geschulte Fachperson ein Finanzierungsbegehren anhand verschiedener Kriterien und aufgrund ihrer Erfahrung. Eine Software hingegen untersucht die Daten anderer Bankkunden, analysiert deren persönliche Angaben und ihre finanzielle Situation und weiss zudem auch, wer von ihnen den Kredit pünktlich bedient. Aus diesen Daten erstellt die Software ein Vorhersagemodell für die Wahrscheinlichkeit, dass eine neue Kundin ihren Kredit abbezahlen wird. Selbstlernende Systeme, die sich durch Training ständig verbessern, werfen ein grundlegendes Problem auf: Wer trägt die Verantwortung für eine Entscheidung, wenn die Prozesskette, die dazu geführt hat, im Dunkeln liegt, weil sie von niemandem mehr nachvollzogen werden kann? Selbst wenn am Ende ein Mensch die Ergebnisse der Software beurteilt und die Entscheidung trifft, wird es heikel, wenn dieser sein Werkzeug immer schlechter versteht.

Verzerrungen, mangelnde Fairness und Vertrauensdefizite

Wenn KI die Wege erarbeitet, um Probleme zu lösen, ergeben sich eine Reihe grundsätzlicher Schwierigkeiten.

Wenn umfangreiche Datensätze der KI als eigentlichen Treibstoff dienen, sitzt diese den Fehlern oder Schlagseiten auf, die sich in den Daten verbergen. Allfällige Einseitigkeiten aber können den Algorithmus prägen – oder ihn gar absichtlich in die Irre leiten, falls entsprechende Lerndaten manipuliert

wurden. Die allfällige Unausgewogenheit grosser Datensätze – die beispielsweise aufgrund von Anfragen in Suchmaschinen zusammengetragen werden – ist indes nicht zwangsläufig auf Arglist zurückzuführen. Vielmehr ist sie unter Umständen historisch gewachsen, indem sie nicht hinterfragte Werturteile und Gewohnheiten widerspiegelt. In diesen Fällen kann KI dazu beitragen, solche Verzerrungen aufzudecken – sofern sie mit der gebotenen Sorgfalt eingesetzt wird.

Ein weiteres Problem ist der zunehmenden Selbstständigkeit der KI geschuldet. Zwar startet die Software bei einem dem Programmierer bekannten Ausgangszustand. Insbesondere bei sogenannten künstlichen neuronalen Netzwerken verändern sich aber im Laufe der zahlreichen Trainingszyklen Stärke und Gewichtung einzelner Verbindungen, sodass die Software am Ende Lösungen vorschlägt, deren logische oder physikalische Basis sich kaum mehr ergründen lässt. Für den praktischen Einsatz sind solche undurchsichtigen KI-Algorithmen, die als eigentliche Blackbox agieren, von beschränktem Nutzen, wenn man verstehen will, wie ein System seine Schlussfolgerungen zieht.

Nicht nur die Daten können Verzerrungen aufweisen, sondern auch die Algorithmen als solche. Denn die Modelle werden von den Entwicklern festgelegt, die dabei unter Umständen gewissen Wertvorstellungen und Interessen den Vorrang gegenüber anderen einräumen. Solche «unfairen» Algorithmen haben bereits zu ungerechten Ergebnissen geführt, indem beispielsweise Menschen aus ärmeren Wohngebieten eine ausgeprägtere Neigung zur Straffälligkeit unterstellt wurde.

Empirische Untersuchungen zeigen, dass viele Personen selbst dann einer menschlichen Entscheidung mehr Vertrauen entgegenbringen, wenn sie wissen, dass die KI erwiesenermassen objektiver entscheidet. Doch es gibt auch Untersuchungen, die auf das Gegenteil hinweisen, indem sie zeigen, dass sich Menschen manchmal allzu stark auf die Ergebnisse automatisierter Entscheidungsprozesse verlassen. Das Mass des Vertrauens richtig einzuschätzen, das die KI verdient, erweist sich als schwierig.

Schliesslich zeichnet sich ab, dass durch KI wirtschaftliche Vormachtstellungen verstärkt werden: Weil zahlreiche neue Formen der KI auf riesige Datensätze angewiesen sind, geniessen Firmen einen Wettbewerbsvorteil, die auf viele Daten zugreifen können. So engagieren sich denn auch führende Tech-Firmen aus den USA und China, die sich die

Angaben ihrer Kundschaft zunutze machen können, in der Entwicklung von KI. Das in der Internetwirtschaft bekannte Phänomen, wonach die ohnehin gut Positionierten rasch stärker werden und dadurch Oligopole entstehen, dürfte sich verschärfen.

Vielfältige Anwendungsfelder

Bereits heute wird KI in den unterschiedlichsten Gebieten angewendet. In der Möglichkeit, Dienste und Angebote passend auf die einzelne Person zuzuschneiden, liegt eine der Stärken von KI-Systemen.

Immer öfter zum Einsatz kommt KI beim Einkaufen, in den klassischen und den neuen Medien, in der Arbeitswelt, in Bildung und Forschung und in Behörden. Diese fünf Anwendungsgebiete nimmt die Studie von TA-SWISS unter die Lupe. Die in dieser Kurzfassung gewählte Reihenfolge weicht von jener im ausführlichen Studienbericht ab. Sie widerspiegelt die alltägliche Erfahrung der meisten Nutzerinnen und Nutzer, die öfter Onlineshopping betreiben und am Bildschirm Nachrichten lesen dürften, als mit Behörden in Kontakt zu treten. Hingegen wird damit nichts über die Bedeutung des entsprechenden Einsatzgebietes ausgesagt.

KI und Konsum: Individualisiert auf Einkaufstour

Wer im Internet Waren bestellt, kommt zwangsläufig mit KI auf Tuchfühlung. Das personalisierte Einkaufserlebnis ist selbstverständlich geworden, die informationstechnischen Prozesse wirken diskret im Hintergrund – und werden kaum hinterfragt.

«Folgende Produkte könnten Sie auch interessieren», «gleich mitbestellen!» oder «passend dazu»: Solchen Formulierungen begegnet, wer im Internet einkauft und dabei von einschlägigen Shops gleich noch auf weitere Produkte hingewiesen wird, die für sie oder ihn reizvoll und praktisch sein könnten. Derartige Empfehlungssysteme nutzen die «digitalen Fussabdrücke» der Kundschaft: Ausgehend von der KI-gestützten Analyse sämtlicher auf der entsprechenden Plattform getätigten Käufe und den persönlichen Angaben der Kunden, aber auch aufgrund ihres sonstigen Verhaltens wie der Verweildauer auf der Seite oder der Mausbewegungen zieht das System Rückschlüsse auf die Vorlieben bestimmter Kundengruppen.

Die Werbung wird ebenfalls zunehmend personalisiert. Auf diesem Feld aktiv sind Google und Grossanbieter wie Alibaba und Amazon, aber auch soziale Netzwerke wie Facebook. Zielgerichtet bewerben sie die unterschiedlichsten Güter und Dienstleistungen, oftmals im Auftrag anderer Firmen. Die im Hintergrund wirkende Technik ist vergleichbar mit jener der Empfehlungssysteme.

Persönliche Vorlieben als Orientierungsmarken

KI-gestützte Empfehlungssysteme bieten der Konsumentin und dem Konsumenten zunächst einmal den Vorteil, dass ihnen die Orientierung im Konsumkosmos leichter fällt. Und statt dass Werbung breit gestreut wird, sollen auf dem Bildschirm nur Anzeigen für Produkte erscheinen, die auf die Interessen der jeweiligen Person zugeschnitten sind.

Was der Kundschaft nützt, um sich in der bunten Warenwelt zurechtzufinden, hilft den Anbietern, ihre Ausgaben zu mindern. Verarbeitungskosten lassen sich dank des Einsatzes von KI um bis zu 20 Prozent senken, und die Werbung fährt weniger Streuverluste ein, wenn sie sich auf die richtige Zielgruppe konzentriert. Auch die Kundenpflege erreicht ein neues Niveau, indem die Ansprache zunehmend individuell erfolgt. Zudem lassen sich die Gewinne maximieren – etwa durch dynamische Preisgestaltung: Dank KI können die Anbieter die Zahlungsbereitschaft ihrer Kundschaft ermitteln und den Warenwert vollautomatisiert entsprechend kalkulieren.

Mensch oder Maschine?

Nicht nur beim Kauf von Gütern, sondern auch für Dienstleistungen wird zunehmend KI verwendet. So stellte Google im Mai 2018 den digitalen persön-

lichen Assistenten Duplex vor, der mit einer täuschend echt klingenden Frauenstimme im Namen seiner Besitzerin einen Friseurtermin vereinbarte oder einen Tisch im Restaurant reservierte. Weil Duplex sogar in den technisch aufgeschlossenen USA vielen Leuten unheimlich vorkam, richtete Google den Assistenten so ein, dass er sich zu Beginn eines Gesprächs selbst vorstellt. In seinem Blog weist der Konzern auf weitere Einsatzmöglichkeiten dieser Technik hin, so etwa für verschiedene Kunden- oder behördliche Auskunftsdienste: Statt lange Warteschlangen in Kauf nehmen zu müssen, bis ein persönlicher Ansprechpartner verfügbar ist, könnten zunehmend digitale Assistenten die häufigsten Fragen beantworten.

Duplex illustriert ein Grundproblem von KI augenfällig: Wer Güter und Dienstleistungen über Internet oder über das Handy ordert, weiss immer seltener, ob eine Person aus Fleisch und Blut Antwort gibt oder ob die Interaktion mit einem Algorithmus stattfindet. Zudem ist für die Kundschaft nicht nur schwer ersichtlich, ob und wann KI-Systeme eingesetzt werden, sondern auch wie und wofür. Selbst wenn Konsumenten darauf hingewiesen werden, dass ihre Gespräche oder ihre persönlichen Daten aufgezeichnet werden, wissen sie nicht, welche Aussagen die Algorithmen später mit welcher Treffsicherheit daraus gewinnen werden.

Digitale Leimruten

Ein Konto, in dem die Informationen über alle bisherigen Bestellungen einer Person und auch ihre individuellen Daten abgelegt sind, ist eine praktische Sache: Das personalisierte Eingangsportal mit Bildern der letzten getätigten Einkäufe vermittelt ein anheimelndes Gefühl, und da Adresse und Kreditkartennummer gespeichert sind, entfallen bei jeder weiteren Bestellung lästige Eingaben. Im Lauf der Zeit – d.h. mit jeder getätigten Transaktion – lernt das System seine Kundschaft besser kennen, sodass die Empfehlungen deren Geschmack genauer treffen. Die Kehrseite der Medaille: Der Komfort verführt dazu, bei einem Anbieter zu bleiben, auch wenn ein anderer möglicherweise ein vorteilhafteres Sortiment anbietet. In der Fachwelt ist dieser Effekt als Plattformklebrigkeit («stickiness») bekannt.

Hinzu kommt, dass auch die schiere Grösse eines Netzwerks die Einzelnen bindet. Denn es ist nützlich, Teil eines möglichst weit gespannten Netzes zu sein, mit dem man viele erreicht. Vorteilhaft ist auch, wenn über eine einzige Plattform verschie-

dene Tätigkeiten erledigt werden können, sodass man sich nicht für jede Transaktion oder Interaktion separat anmelden muss. Die Folge: Am Ende konzentrieren sich die Daten bei einigen wenigen Grosskonzernen mit erheblicher Marktmacht.

Schutz der Daten schützt auch die Person

Unsere Spuren im Internet sind aufschlussreich. Das soziale Netzwerk Facebook selbst veröffentlichte eine Studie mit über 8'000 freiwilligen Usern, deren persönliche Merkmale ein KI-System anhand getätigter «Likes» treffender eingeschätzt hatte als die Freunde der User. Nur zehn Likes brauchte die Maschine, um bessere Ergebnisse zu liefern als Mitarbeiter; ab siebzig analysierten Likes übertraf die maschinelle Beurteilung jene aus dem Freundeskreis. Andere Studien zeigten, dass KI-Systeme in der Lage sind, aus dem Profilbild eines Menschen seine individuellen Eigenschaften – einschliesslich der sexuellen Orientierung – erstaunlich genau herauszulesen.

Ihr Leistungsvermögen vermag KI insbesondere dann auszuspielen, wenn sie persönliche Daten aus verschiedenen Quellen kombiniert. Das ist ein Vorteil, den Firmen wie Facebook und Google zu nutzen wissen, denn ihnen gehören mittlerweile zahlreiche beliebte Dienste: Facebook besitzt unter anderem den Fotodienst Instagram und den Messenger WhatsApp, während Google – nebst vielen anderen Firmen – das Videoportal YouTube und das Fitnesstracker-Unternehmen Fitbit sein Eigen nennt. Auch andere Techgiganten wie Amazon und Alibaba haben eine Vielzahl von Web-Plattformen und Datensätzen gekauft, sodass sie Daten verschiedenen Ursprungs miteinander verknüpfen und damit ihren Wettbewerbsvorteil gegenüber kleineren Unternehmungen ausbauen können. Nicht alle Fachleute erkennen indes in den konzentrierten Datensätzen eine Gefahr; denn die User können die Daten mehreren Anbietern zur Verfügung stellen, sodass der Zugriff auf bestimmte Informationen als solcher noch keinen Vorsprung gegenüber Rivalen garantiert.

Der Kundschaft die alleinige Sorge um den Schutz der Daten aufzubürden, ist keine Lösung. Zwar müssen die Allgemeinen Geschäftsbedingungen (AGB) Aufschluss darüber geben, welche Angaben erhoben und gesammelt werden. Doch selbst wenn jemand den AGB zustimmt, kann daraus noch nicht abgeleitet werden, dass die jeweilige Person damit den Zugriff auf ihre Daten gestattet, weil ihr der Umfang

der erhobenen Informationen und die Nutzung derselben kaum je bewusst sein dürfte: Allein um die AGB seiner am häufigsten genutzten Dienste vollständig zu lesen, müsste ein durchschnittlicher User jährlich mehrere Wochen aufwenden.

Schwierige Balance zwischen Personalisierung und Datenschutz

Dass unter den Nutzerinnen und Nutzern das Bewusstsein für den Wert der eigenen Daten zunimmt, ist aus Sicht der Fachleute, die im Rahmen der TA-Studie befragt wurden, ein vordringliches Anliegen. Für Unternehmen ist es allerdings schwierig, Personalisierung und Datenschutz miteinander in Einklang zu bringen. So jedenfalls lautet die Einschätzung der befragten Expertinnen und Experten. Auch drohende Vertrauensverluste sind Probleme, denen Anbieter im Web entgegenzutreten sollten. Offene

Daten mit Kontrollmöglichkeiten für Kundinnen und Kunden, transparentes Offenlegen des KI-Einsatzes sowie das Recht auf Datenlöschung werden von den befragten Fachpersonen als aussichtsreiche Massnahmen gesehen, um das Kundenvertrauen zu fördern.

Die Entstehung von Oligopolen, in denen ein paar wenige Giganten den Markt beherrschen, wird von den Befragten als weitere erhebliche Gefahr genannt, die mit KI-Systemen einhergehen dürfte. Denn letztlich könnten solche Strukturen zu einer Schmälerung des Angebots führen und die Wechselmöglichkeiten der Kundschaft einschränken. Entsprechend wichtig ist es sicherzustellen, dass diese ihre personenbezogenen Daten beim Wechsel des Anbieters mitnehmen kann. Auch gilt es, Verschiebungen im Markt genau zu beobachten und über Kontroll- und Eingriffsmöglichkeiten zu verfügen, die den Wettbewerb sicherstellen.

KI und öffentliche Kommunikation: Meinungsmache im medialen Echoraum

Innerhalb der letzten zehn Jahre haben Presse, Radio und Fernsehen an Bedeutung eingebüsst: Immer mehr Menschen informieren sich über soziale Medien wie Facebook oder Twitter. Professionell recherchierte Nachrichten, die in einen breiteren Zusammenhang eingeordnet werden, müssen sich gegen Blogbeiträge oder Posts aus sozialen Netzwerken behaupten.

Im Jahr 2017 bezogen 83 Prozent der Bevölkerung in der Schweiz ihre Nachrichten online. 45 Prozent nutzten dabei die sozialen Medien, wobei nebst Facebook und Twitter immer stärker auch Plattformen wie Google News und Apple News herangezogen werden, die Inhalte aus anderen Quellen bündeln. Die Folge: Mehr und mehr User teilen Artikel und informieren sich direkt untereinander und teilen dabei oft auch Nachrichten traditioneller Anbieter. Der Ursprung der weitverbreiteten News wird dabei zunehmend unklar.

Bestärkt in der Blase

Zum Anspruch herkömmlicher Medien gehört es, ausgewogen über eine Fülle an Themen zu informieren und dabei die herrschende Meinungsvielfalt abzubilden. Wer sich in der Presse oder über Radio und Fernsehen informiert, erfährt zwangsläufig auch etwas über Sachverhalte, die sie oder ihn weniger interessieren.

Anders verhält es sich, wenn News über soziale Netzwerke geteilt werden. Treffen auf dem Konto einer Person viele Newsfeeds ein, bestimmt der Algorithmus, welche Inhalte zuerst gezeigt werden – und zwar ausgehend von den gesammelten Informationen über den Kontoinhaber, die dieser entweder freiwillig offengelegt oder über sein bisheriges Onlineverhalten preisgegeben hat. Dabei neigen KI-Systeme dazu, den Nutzern Information vorzuschlagen, die deren inhaltlichen Vorlieben entspricht; die User orientieren sich dann vornehmlich in ihrer Filterblase. Wenn zudem Personen nur noch Artikel lesen, die ihnen ihr Freundeskreis vorschlägt, verstärkt sich die Gefahr, dass sie sich in einer Echokammer wiederfinden: Hier

begegnen sie ausschliesslich Inhalten, die sich mit ihren Interessen decken und ihre eigene Meinung widerspiegeln, während abweichende Haltungen und gegenläufige Informationen ausgeblendet werden.

Wenngleich digitale Filterblasen und Echokammern auf erhebliche öffentliche Aufmerksamkeit stossen und als Gefahr für die Demokratie diskutiert werden, gibt es bis jetzt keinen wissenschaftlich erhärteten Nachweis, wonach die Meinungsvielfalt online stärker eingeschränkt wird als in herkömmlichen Meinungszirkeln Gleichgesinnter. Allerdings stehen die geschilderten Entwicklungen noch in ihren Anfängen und könnten sich sehr wohl zu einem Problem für die Meinungsbildung in demokratischen Gesellschaften auswachsen.

Verschörungstheorien aus der Trollfabrik

Fehlerhafte Medienberichte sind kein neues Phänomen und bereichern seit dem frühen 19. Jahrhundert die metaphorische Tierwelt um die Gattung der «Zeitungsenten». Diese tritt meistens unfreiwillig auf und beschädigt den Ruf des Blattes, welches sie veröffentlicht hat. Neu ist hingegen die beabsichtigte massenhafte Fehlinformation über soziale Medien.

Schon heute werden Fake News weitgehend automatisch und sehr effizient verbreitet. Zum Einsatz kommen dabei Social Bots – einfache Programme, die geschrieben werden, um auf sozialen Medien gezielt Nachrichten zu verbreiten. Während des US-Wahlkampfes von 2016 zwischen Hillary Clinton und Donald Trump stammte rund ein Fünftel der abgesetzten Tweets von Social Bots; von den automatisierten Twitter-Konten warb die überwiegende Mehrheit (d.h. bis zu drei Viertel) für Trump.

Künftig könnte KI in der Lage sein, sogenannte Trolle zu imitieren. Bei diesen handelt es sich um User, die durch emotionale und oft mit Falschinformationen angereicherte Blog- und Forumsbeiträgen versuchen, ebenso heftige Antworten zu provozieren. Mit dieser Strategie zielen die oft in eigentlichen «Fabriken» angestellten Trolle darauf ab, die Bekanntheit gewisser Gruppen zu steigern oder gar Verschwörungstheorien zu streuen. Social Bots wie auch Trolle beziehen sich gerne auf Websites, die eigens für die Verbreitung von Fehlinformation angelegt sind und die so die Glaubwürdigkeit ihrer digitalen Handlanger erhöhen. Schon heute potenzieren KI-Systeme die Effizienz, mit der die Trolle ihre Botschaften kolportieren.

Demokratie in Gefahr

Eine Strategie der Absender von Fake News besteht darin, diese in die etablierten Nachrichtenkanäle einsickern zu lassen. In den USA zeigte sich, dass Fehlinformationen zunächst gerne an lokale Medien herangetragen werden, da diese über weniger Möglichkeiten verfügen, um Fakten nachzuprüfen. Einmal von den Lokalnachrichten aufgenommen, gelangen die Fake News rasch in die nationalen Medien. KI hilft aber nicht nur beim Verbreiten von Falschinformationen, sondern trägt auch dazu bei, dass diese selbst immer ausgereifter werden. So kann KI eingesetzt werden, um Bilder, Audiodateien und sogar Videos zu manipulieren oder von Grund auf fiktives Filmmaterial herzustellen, das dazu dient, Personen zu diskreditieren oder zu erpressen. Die Fachwelt bezeichnet diese technisch erzeugten unechten Videos als «deep fakes».

Unter autoritären Regimes, die unliebsame Berichterstattung abwürgen und die Medien kontrollieren, eröffnen die sozialen Medien mitunter einen letzten Ausweg für unabhängige Information. In funktionierenden Demokratien hingegen wird die Beeinflussung der öffentlichen Meinung über Fake News in den sozialen Medien als ernste Gefahr für das politische System gesehen. Dies zum einen, weil die Öffentlichkeit das Vertrauen in die etablierten Nachrichtenkanäle verliert: Neueren Umfragen zufolge vertrauen in fast allen Ländern nur noch weniger als die Hälfte der Befragten der Fähigkeit herkömmlicher Medien, Fakten von Fiktion unterscheiden zu können. Zum andern zeigten verschiedene prominente Analysen, dass Falschnachrichten sich sogar besser verbreiten als korrekte Information. Für Deutschland etwa liess sich nachweisen, dass von den zehn über Angela Merkel am weitesten propagierten Nachrichten sieben als Fake News einzustufen sind.

Skeptischer Blick auf soziale Netzwerke

Die Umfragen, die im Rahmen der Studie von TA-SWISS durchgeführt wurden, brachten eine skeptische Haltung gegenüber Social Media zum Ausdruck: Mehrheitlich wird befürchtet, die sozialen Netzwerke seien schädlich für die inhaltliche Qualität des Journalismus. Etwas weniger negativ wird die blosser Onlineverbreitung von Nachrichten eingeschätzt. Die Personalisierung von Inhalten findet unter den Fachleuten auch Befürworter; ausschlaggebend scheint zu sein, dass die User selber entscheiden können, ob, auf welche Weise und wie stark der Nachrichtenfluss auf sie persönlich ausgerichtet wird.

Fake News hingegen werden von den Befragten als ernste Gefahr beurteilt. Nur eine Minderheit ist der Ansicht, dass der technische Fortschritt dazu führen wird, Falschinformationen besser zu entlarven. Die Mehrheit befürchtet im Gegenteil, dass ausgereifere Technik zunehmend raffiniertere Fake News erzeugen wird. Die Umfrageresultate förderten allerdings auch zutage, dass es im Kampf gegen Falschinforma-

tion das Recht auf freie Meinungsäußerung im Blick zu behalten gilt. Dass für die Prüfung von Inhalten eigens Behörden eingerichtet werden, wird eher kritisch beurteilt. Hingegen begrüßen es die Befragten, wenn die Plattformbetreiber Fake News löschen und deren Urheber sperren. Das automatisierte – d.h. KI-gestützte – Durchchecken von Fakten wird dabei als hilfreicher eingestuft als manuelle Prüfverfahren.

KI in der Berufswelt: Wenn Maschinen Kopfarbeit verrichten

Lange galten nur einfache und repetitive Tätigkeiten als gefährdet, der Rationalisierung zum Opfer zu fallen. Seit KI in Berufsfelder vordringt, die als anspruchsvoll gelten und eine gute Ausbildung voraussetzen, geht auch in der Teppichetage die Furcht vor Arbeitsplatzabbau um.

Um die Jahrtausendwende erheiterten verschiedene Zeitungskolumnen die Leserschaft mit maschinellen Übersetzungen kurzer Texte. Die Programme scheiterten an den einfachsten Sprachbildern und lieferten eigenwillige und oft komische Traktate. Bis vor Kurzem galt die Übersetzung als eine Aufgabe, die von einer Maschine nicht zu bewältigen sei. Doch Software wie DeepL, die mit künstlichen neuronalen Netzen arbeitet, liefert heute durchaus brauchbare Übersetzungen, während andere Programme gar eigenständig einfache journalistische Texte erzeugen können.

Längst sind also nicht mehr nur einfache Routine-tätigkeiten von der Automatisierung betroffen: Im Einsatz ist KI in der Medizin, wo sie bei der Beurteilung von Ultraschall-, Magnetresonanz- oder Röntgenbildern treffendere Ergebnisse liefert als die meisten medizinischen Fachleute, oder auch in der Rechtsprechung, wenn sie Juristen dabei unterstützt, Tausende von Vertragsseiten nach bestimmten Klauseln zu durchsuchen.

Zahlen zum drohenden Arbeitsplatzverlust – oder, je nach Betrachtungsweise, zum Rationalisierungspotenzial – liefern eine Reihe von Studien. Die Spannweite der Schätzungen ist breit: Genannt werden mögliche Jobeinbussen zwischen 6 bis 57 Prozent, je nach Grundannahmen und angewandter Methode. Einig ist sich die Fachwelt darin, dass es – anders

als bei früheren innovationsbedingten Umbrüchen der Wirtschaft – schwierig sein dürfte, den durch KI bedingten Beschäftigungsschwund durch neue Produkte und Dienstleistungen wettzumachen.

Karriere unter der Schirmherrschaft von KI

KI wirkt indes nicht nur auf die Quantität, sondern auch auf die Qualität der Arbeit zurück. So setzen Personalverantwortliche bei der Rekrutierung von Arbeitskräften zunehmend auf KI-Systeme, um die Bewerbungsdossiers vorzuselektionieren. Die Bewerberinnen und Bewerber müssen hierfür zunächst online standardisierte Fragebögen ausfüllen. Anschliessend arbeiten die Personalverantwortlichen aufgrund der KI-erstellten engeren Auswahl weiter. Für die Kandidaten muss das nicht negativ sein: Wenn von Anfang an ein Mensch die eingereichten Unterlagen prüft, können unsachliche Entscheidungskriterien wie das fotogene Aussehen einer Person übermässiges Gewicht erhalten. Dagegen birgt der Einsatz eines KI-Systems die Chance, bestehende diskriminierende Praktiken offenzulegen und zu ändern.

Grosskonzerne setzen vermehrt KI ein, um Kompetenzprofile der einzelnen Angestellten zu erstellen und auf das Individuum abgestimmte Weiterbildungen vorzuschlagen. Heikel ist, dass dabei die Autonomie der Einzelnen beschnitten wird, weil nicht sie sich für weitere Qualifizierungsmassnahmen entscheiden, sondern ein System versucht, die vorhandenen menschlichen (und finanziellen) Ressourcen zu optimieren. Auch wenn letztlich nicht das System die Entscheidung

trifft, können seine Ratschläge eine Eigendynamik im Betrieb entwickeln. Für die Angestellten wird es schwierig zu begründen, wieso sie einer Empfehlung nicht Folge leisten möchten. Generell wirft der Schutz der persönlichen Daten ein grundsätzliches Problem beim Einsatz von KI in der Arbeitswelt auf.

Arbeit im maschinellen Takt

Künftig werden aufgrund von KI nicht unbedingt ganze Arbeitsstellen als vielmehr einzelne Tätigkeiten wegfallen. Die Arbeitnehmerinnen und Arbeitnehmer werden lernen müssen, mit «klugen» Maschinen zu kooperieren und mit dem gesteigerten Arbeitstempo und der Verdichtung der Tätigkeiten umzugehen. Fachleute und Institutionen wie die internationale Arbeitsorganisation ILO betonen die Notwendigkeit, die Menschen bei der Gestaltung der Arbeit in den Mittelpunkt zu stellen, damit sie durch den Einsatz von KI-Systemen nicht schlechter gestellt werden als bisher. Auch wird es vor allem bei nicht erklärba- ren Resultaten der KI wichtig sein, dass die Angestellten in der Lage sind, die Plausibilität der Ergebnisse einzuschätzen und grobe Fehler zu verhindern.

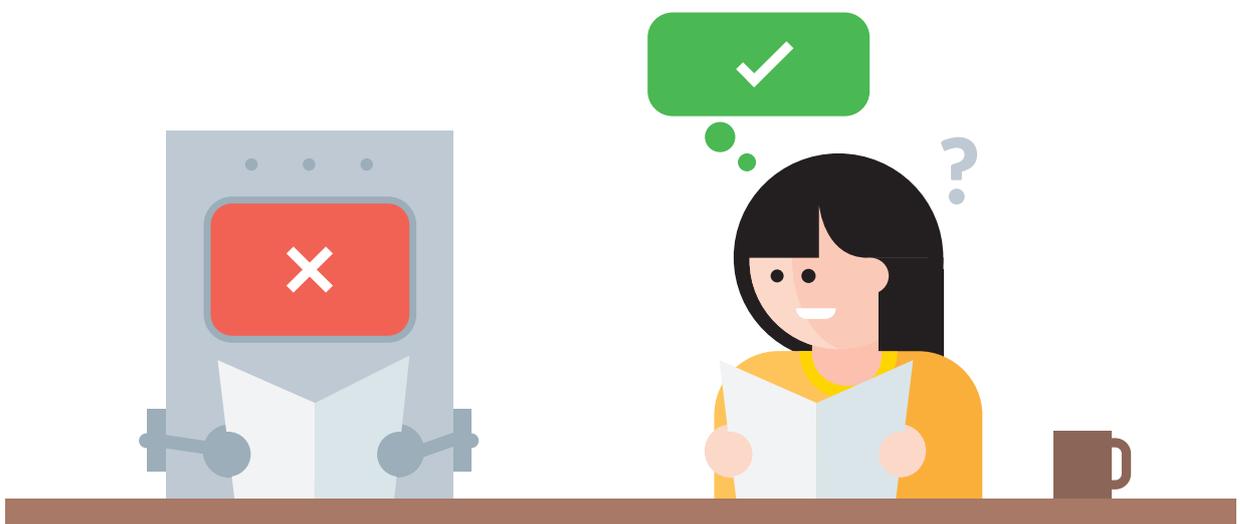
Der Einfluss von KI auf die Arbeitsgestaltung lässt sich nur schwer von den Effekten der Digitalisierung im Allgemeinen trennen. Wie die Studie von TA-SWISS über flexible Arbeit gezeigt hat, kann die Digitalisierung dazu beitragen, die persönliche Freiheit und Zeitautonomie zu erhöhen. Sie stellt allerdings auch das bisherige System sozialer Absicherung in Frage. Im Netz existiert eine Vielzahl von Plattformen, die Grossaufträge stückeln und

portioniert an sogenannte «Clickworker» vermitteln. Die Studentin oder der Student kann so nebenher am Abend ein Taschengeld dazuverdienen – doch insbesondere für selbständige Kreative, die auf solchen Plattformen ihre Aufträge im internationalen Wettbewerb akquirieren, erhöhen sich Druck und existenzielle Unsicherheit.

Massnahmen gegen die Zweiteilung des Arbeitsmarktes sind gefragt

Die im Rahmen der Studie von TA-SWISS befragten Fachleute halten es für wahrscheinlich, dass KI die Zweiteilung des Arbeitsmarktes vorantreiben wird. Sie erwarten, dass sich die Schere zwischen Arbeitnehmern mit hoher Qualifikation und entsprechenden Löhnen und solchen mit tiefer Ausbildung und einem geringen Einkommen weiter öffnen wird. Auch sehen die Experten eine stärkere Kontrolle der Angestellten sowie eine wachsende Instabilität der Arbeitsverhältnisse voraus. Dass KI zum Rückgang der Arbeitszeit und -belastung beitragen wird, wird dagegen als unwahrscheinlich erachtet. Alles in allem hält es die Mehrheit der Experten für wahrscheinlich, dass Grossunternehmen besonders stark von diesen Veränderungen in der Arbeitswelt profitieren werden.

In der Weiterbildung erkennen die befragten Fachleute eine besonders wirkungsvolle Massnahme gegen Polarisierungseffekte. Die Frage nach geeigneten Vorkehrungen gegen die wachsenden Gegensätze in der Berufswelt scheint ebenso sehr zu polarisieren wie die KI den Arbeitsmarkt. So wird die Abfederung sozialer Ungleichheiten durch einen entsprechenden



Ausbau des Sozialsystems von einer vergleichsweise geringen Mehrheit der Befragten befürwortet; eine ebenso knappe Mehrheit lehnt es ab, die Arbeitszeit herabzusetzen, um die abnehmende Anzahl Stellen auf mehr Menschen zu verteilen. Die Einführung eines bedingungslosen Grundeinkommens wird wiederum von einer relativen Mehrheit befürwortet. Die Studienautoren selbst erachten hingegen eine bessere und flexiblere Verteilung von Arbeit und Einkommen als kurz- bis mittelfristig zu bevorzugende Alternative. Auch sollten neuartige Beschäftigungsmöglichkeiten wie etwa kreative Startups gefördert und innovative Formen der Besteuerung entwickelt werden, um Staatshaushalt und Sozialsystem weiter finanzieren zu können.

Ein Anliegen ist es den Befragten auch, die Angestellten vor einer ausufernden Kontrolle durch ihre Arbeitgeber zu schützen. Massnahmen gegen die Überwachung am Arbeitsplatz wie auch generell ein besserer Datenschutz werden als wünschenswert und wirkungsvoll eingeschätzt, wobei Arbeitnehmende auch in Arbeitsverhältnissen zu schützen sind, die über staatliche Grenzen hinausgreifen. Verschiedene Fachleute heben indes auch die positiven Potenziale der KI hervor, sofern diese eingesetzt wird, um monotone Arbeiten zu automatisieren und Produktivität wie auch Wettbewerbsfähigkeit zu erhöhen. Würden die erreichbaren Vorteile gerecht verteilt, nütze das der ganzen Gesellschaft.

KI in Bildung und Forschung: Unterstützung für die Humanintelligenz

KI ist ein Produkt der Forschung – und sie hilft ihrerseits, die Wissenschaft weiter voranzutreiben. In der Schule können KI-Systeme dazu dienen, Schülerinnen und Schüler individuell zu fördern.

Bereits jetzt steht die individuelle Förderung von Schülerinnen und Schülern auf der bildungspolitischen Agenda. KI stellt Möglichkeiten in Aussicht, dieses Anliegen technisch zu unterstützen.

Jedem Kind seine Superlehrkraft

Zahlreiche Start-Ups haben bereits Werkzeuge entwickelt, um eine Schülerin oder einen Schüler an ein Ausbildungsziel heranzuführen. Zum Einsatz kommen beispielsweise Chatbots: Programme, die automatisch auf Fragen und Mitteilungen der Lernenden reagieren, indem sie aufgrund von Schlüsselwörtern und Regeln vorprogrammierte oder durch das KI-System individuell erzeugte Antworten geben.

Andere Programme führen die jungen Menschen Schritt für Schritt durch Tutorials, die ihnen die gewünschten Kenntnisse in Mathematik, Naturwissenschaft oder Physik beibringen. Diese KI-Systeme tun dies gestützt auf eingangs ermittelte Kompetenz- und Interessensprofile, anhand derer sie Stärken, Schwächen und bestehende Vorkenntnisse des

Kindes ermitteln; in der Folge analysieren sie seinen Lernweg und schlagen ihm den am besten passenden nächsten Lernschritt vor.

Ein weiterer Ansatz besteht darin, das KI-System den Schreibprozess von Lernenden überwachen und laufend Rückmeldungen geben zu lassen. Auch existieren KI-Systeme, die das Lernen in Gruppen organisieren, indem sie Schülerinnen und Schüler mit zueinander passenden Profilen zusammenbringen oder Jugendliche, die an Mathematik scheitern, mit Mathe-Tutoren aus anderen Klassen, Schulen oder gar Ländern in Kontakt setzen.

Den Lehrkräften geht KI ebenfalls zur Hand – insbesondere bei der Leistungsbeurteilung ihrer Schützlinge, aber auch, um diese individuell zu fördern. Allein bei ihren administrativen Aufgaben können Lehrpersonen dank des Einsatzes von KI wöchentlich etliche Arbeitsstunden einsparen.

Auf der Lauer nach den Spitzenkräften der Zukunft

Wenn Schülerinnen und Schüler ihre Aufgaben zunehmend online im Dialog mit einem KI-System lösen, spült das eine Flut persönlicher Daten auf die Server seines Anbieters. Dieser erfährt nicht nur

vieles über die Lernstrategien und Fähigkeiten der jungen Menschen, sondern auch über ihre Vorlieben und Freizeitaktivitäten. Solches Wissen ist interessant, lassen sich so doch frühzeitig potenzielle Spitzenkräfte ausmachen. Auch die Daten leistungsschwächerer Kinder sind nützlich, denn diese sind womöglich empfänglich für einen Nachhilfekurs, den die gleiche Firma anbietet.

Setzt eine Schweizer Schule KI-Software von privaten Anbietern ein, stellen sich eine Reihe datenschutzrechtlicher Probleme in besonderer Schärfe: Wer darf welche persönlichen Daten für wie lange speichern, besitzen oder abfragen, und wie ist es zu gewährleisten, dass sie gegebenenfalls auch gelöscht werden können? Ist es möglich, die eigenen Daten bei einem Schulwechsel mitzunehmen – und dürfen andere Schulen die Daten einsehen, wenn sie neue Schüler aufnehmen? Wie kann sichergestellt werden, dass solche Prüfungs- und Lerndaten künftig den Wettbewerb um Arbeitsstellen nicht verzerren, wenn die Einen günstige Resultate vorweisen können, die Anderen aber über gar keine Daten oder nur über unvorteilhafte verfügen? Um die wesentlichen Risiken eines problematischen Umgangs mit persönlichen Daten zu beheben, hat die Schweizerische Konferenz der kantonalen Erziehungsdirektoren EDK die Plattform Edulog.ch lanciert. Diese gestattet es Schweizer Schülern, sich anonymisiert in Bildungsprogramme einzuloggen. Die Kantone machen freiwillig mit, und es wird sich noch weisen müssen, ob das Angebot in der Praxis genutzt wird.

Wachsender Einfluss von Privatunternehmen auf die Bildung

Mehrheitlich begrüßen die im Zuge der Studie von TA-SWISS befragten Fachleute die Möglichkeit, Lehrinhalte zugeschnitten auf den einzelnen Schüler zu vermitteln. Denn diese Individualisierung ermöglicht es, auf spezifische Probleme jedes Lernenden einzugehen. Sorge bereitet den befragten Experten allerdings, dass Privatunternehmen mehr und mehr Einfluss auf die Schulen nehmen, die dadurch in Abhängigkeit eines Anbieters geraten könnten.

Ein Verbot, private KI-Systeme zu nutzen, wird allerdings nicht als gangbarer Weg gesehen. Vielmehr sollten mit den Anbietern präzise und verbindliche Vereinbarungen ausgehandelt werden, die auch klare Regelungen zum Datenschutz umfassen. Ob die Entwicklung schulischer KI-Systeme den Behörden überantwortet werden sollte, beurteilen die

Fachleute uneinheitlich; die Zusammenarbeit zwischen Verwaltung und privaten Anbietern wird indes für äusserst wünschenswert gehalten. Mehrheitlich wird auch eine behördliche Prüfung schulischer KI-Anwendungen empfohlen.

KI wird erforscht – und bringt Forschung weiter

In absoluten Zahlen gemessen erscheinen die meisten Publikationen zu KI in China und den USA. Doch auch die Forschungsaktivitäten der Schweiz lassen sich sehen: Umgerechnet auf die Bevölkerungszahl werden hierzulande mehr Wissenschaftsartikel zu KI veröffentlicht als in Deutschland oder Frankreich; gewichtet nach den Auswirkungen der Zitate, steht die Schweiz gar an der Weltspitze.

KI kommt dabei in den unterschiedlichsten wissenschaftlichen Disziplinen vornehmlich dann zum Einsatz, wenn riesige Datenmengen analysiert werden müssen; so wird beispielsweise in der Materialwissenschaft, aber auch in der Biochemie oder in der Astrophysik auf KI zurückgegriffen. In der Lehre ist diese ebenfalls von Nutzen – etwa wenn es darum geht, Schummeleien in wissenschaftlichen Arbeiten aufzudecken. KI-gestützte Werkzeuge spüren sowohl ähnlich klingende Textpassagen in Plagiaten als auch gefälschte Statistiken auf. Ebenso vermag die KI einen Wechsel des Schreibstils, der auf Ghostwriting hindeutet, zu detektieren.

Schlüsselkompetenzen ausbauen und Transparenz sicherstellen

Für die Forschung gewinnt KI laufend an Bedeutung; für grössere Entwicklungssprünge ist sie oft gar eine Voraussetzung. Die befragten Experten sind daher mehrheitlich der Ansicht, wissenschaftliche Einrichtungen sollten Anlaufstellen für Weiterbildung und Nutzung von KI in der Forschung einrichten, die auch als Knotenpunkt für die interdisziplinäre Auseinandersetzung mit diesen Techniken dienen könnten.

Mit Blick auf die Daten, die durch KI erzeugt werden, ist sich die Mehrheit der Befragten einig, dass diese Eigentum der jeweiligen Forschungseinrichtung bleiben oder im open access-Format allen zugänglich gemacht werden müssen. Das mit KI-Systemen erzeugte Wissen schliesslich muss so weit möglich transparent und reproduzierbar sein. Um diese Forderung zu erfüllen, braucht es entsprechende Transparenzrichtlinien.

KI in der Verwaltung: Hoheitliche Befugnisse für künstliche Intelligenz?

Behörden arbeiten in der Regel mit Formularen und nach standardisierten Vorgehensweisen. Ideale Voraussetzungen also für den Einsatz von KI. Der gesteigerten Verwaltungseffizienz steht das Risiko gegenüber, dass Grundrechte der Bürgerinnen und Bürger beschnitten werden.

Im September 2018 reichte ein Schweizer Nationalrat der sozialdemokratischen Partei eine Motion ein, die den Bundesrat mit der Einrichtung eines Kompetenzzentrums in der Verwaltung beauftragte. Dieses solle den Einsatz von KI und maschinellem Lernen unterstützen, um damit die Effizienz der Behörden zu steigern. Der Bundesrat hat empfohlen, die Motion abzulehnen.

Im Einsatz gegen Betrug und Verbrechen

KI ist bis in die Amtsstuben vorgedrungen. So bedienen sich beispielsweise australische Steuerbehörden der KI, um Geldforderungen des Staates automatisch einzutreiben. Auf Kritik stiess das System, weil es teilweise inkorrekte Schlüsse zog und es den Steuerpflichtigen überlassen blieb, sich gegen zu hohe Abgaben zu wehren. In Deutschlands Steuerämtern wird KI verwendet, um Schummeleien in den Steuererklärungen aufzudecken.

Wenngleich in der Schweiz der Einsatz von KI im Besteuerungsverfahren (noch) nicht diskutiert wird, schreitet auch hier die Digitalisierung voran: Zölle werden zunehmend digital erfasst und verarbeitet, auch verschiedene Verbrauchssteuern – etwa auf Alkohol, Mineralöl und Tabak – könnten künftig automatisiert erhoben werden. Entsprechende Rechtsgrundlagen sind im Anhang des im September 2015 verabschiedeten Entwurfs zum neuen Datenschutzgesetz vorgesehen.

Von der Polizei wird KI zum einen genutzt um abzuschätzen, in welchem Quartier sich demnächst ein Verbrechen ereignen könnte. Dank solcher Voraussagen sollen die Einsatzkräfte verstärkt dort patrouillieren können, wo die Wahrscheinlichkeit für einen Einbruch oder einen Überfall besonders hoch ist. Vorausschauende Polizeiarbeit ist vor

allem in einigen US-amerikanischen Grosstädten gebräuchlich. Ein anderer Ansatz fokussiert nicht auf den Raum, sondern auf die Personen und ihr Beziehungsnetz: Mit den Daten von Sozialkontakten in bestimmten Kreisen errechnet das KI-System die Wahrscheinlichkeit, dass jemand Mitglied einer kriminellen Bande ist.

Des Weiteren hilft KI bei der Einschätzung der Rückfallgefahr von Straftätern. In den USA ist ein KI-System im Einsatz, das aufgrund von 137 Merkmalen die Wahrscheinlichkeit errechnet, dass ein Gesetzesbrecher erneut straffällig wird. In die Kritik geriet das System, weil es farbigen Straftätern systematisch eine höhere Rückfallgefahr unterstellte als weissen. Zwar wird auch in der Schweiz Software bei der kriminalistischen Beurteilung eingesetzt – insbesondere um zu ermitteln, ob bei einer bestimmten Person der Strafvollzug gelockert werden kann. Diese Überprüfung beruht aber nicht auf fortgeschrittenen KI-Technologien.

Grundrechte weisen staatliche Willkür in die Schranken

Bei der Erfüllung ihrer hoheitlichen Aufgaben verarbeiten Behörden eine grosse Anzahl an Daten von Personen. Der Einsatz von KI-Systemen verspricht hier eine erhebliche Effizienzsteigerung: Dokumente und Akten könnten künftig automatisch erstellt werden, während sprachbasierte digitale Büroassistenten Routineanfragen aus der Bevölkerung beantworten würden.

Auf der anderen Seite aber muss die Obrigkeit besondere Sorgfalt bei all ihren Handlungen walten lassen und darf diese nicht allein der Effizienz unterordnen. Die Grundrechte, die dem Individuum gegenüber dem Staat garantiert werden, statuieren, dass der Einzelne vor allfälligen, in grundrechtlich geschützte Positionen eingreifende Akte zu schützen ist.

Der Schutz der persönlichen Daten ist dabei ein grundlegender Anspruch; diese müssen sicher gespeichert und vor Missbrauch geschützt sein. Auch sollen alle Personen davon ausgehen können, dass Daten, die über sie gesammelt und bearbeitet

werden, zutreffend, vollständig und ausgewogen sind. Tatsächlich sind sich aber viele Personen nicht im Klaren darüber, welche ihrer Daten überhaupt gespeichert und im Rahmen einer KI-gestützten Verarbeitung genutzt werden; mithin sind sie auch kaum in der Lage, allfälligen Fehlern in ihrem gesammelten Datensatz nachzugehen und diese zu berichtigen.

Angesichts der Verzerrungen und Unausgewogenheiten, die mit der KI-gestützten Datenverarbeitung einhergehen können, sind insbesondere automatisierte Verfahren heikel. Denn diese könnten zur Benachteiligung bestimmter Personengruppen und Individuen führen. Das revidierte Datenschutzgesetz sieht hier vor, dass die Betroffenen bei vollständig automatisierten Entscheidungen informiert werden müssen.

Einem vollautomatisierten KI-System die Entscheidung zu überlassen, widerspricht schliesslich der rechtsstaatlichen Forderung, wonach eine behördliche Weisung oder ein Urteil zu begründen ist. Jeder amtliche Rechtsakt sollte transparent dargelegt werden, damit sich die Betroffenen allenfalls dagegen zur Wehr setzen können. Wenn aber eine KI-basierte, selbstlernende Maschine automatisch eine Entscheidung trifft oder auch wesentliche Vorarbeiten dazu erledigt, ist es unter Umständen nicht mehr möglich, die Herleitungen des KI-Systems nachzuvollziehen. Ebenso verunmöglicht ist dadurch eine Begründung des Entscheids, die rechtsstaatlichen Ansprüchen genügt.

Zurückhaltende Schweiz

In der Befragung, die im Rahmen der Studie von TA-SWISS stattfand, waren die Fachleute überwiegend der Ansicht, der Einsatz von KI-Systemen in der öffentlichen Verwaltung sei in der Schweiz weniger wahrscheinlich als in anderen Ländern, und eine völlige Automatisierung komplexer behördlicher Abläufe sei kaum zu erwarten. Hierzulande könnten KI-Systeme allenfalls bei einfachen Fragen und in teilautomatisierten Prozessen Verwendung finden.

Hingegen erachtet eine Mehrheit der Befragten das Risiko für beträchtlich, dass Schweizer Behörden in die Abhängigkeit von Technologieanbietern aus dem Ausland geraten könnten. Auch droht aus Sicht der Fachleute eine gewisse Maschinenhörigkeit in dem Sinne, dass Verwaltungsangestellte Mühe bekunden könnten, eine von der KI abweichende Position zu begründen. Schliesslich warnen die befragten Expertinnen und Experten vor einer Schwächung des Datenschutzes und davor, dass automatisierte Abläufe nur mangelhaft kontrolliert werden können.

Keine Delegation von Entscheidungsgewalt an die Maschine

Eine deutliche Mehrheit der befragten Fachleute ist sich einig: Wenn es um vorausschauende Analy-



sen oder die Erstellung eines Persönlichkeitsbildes («Profiling») geht, dürfen KI-Systeme nur Empfehlungen abgeben; die Entscheidung selbst muss den Menschen vorbehalten bleiben. Eine knappe Mehrheit der Befragten ist zudem der Ansicht, KI dürfe nur verwendet werden, um raum- oder objektbezogene Entscheidungen vorzubereiten, nicht aber, wenn es um Beschlüsse geht, die Personen betreffen.

Mehrheitlich begrüsst wird, dass der Staat verpflichtet wird offenzulegen, in welchen Verfahren er KI einsetzt. Ausserdem sollen die Behörden die Grundzüge des Prozesses, für den KI eingesetzt wird, verständlich erklären. Und schliesslich soll der Staat jede Entscheidung, die sich auf KI abstützt, individuell begründen.

«Kluge» Maschinen – eine Herausforderung für geltendes Recht

Wenn ein Ingenieur die Tragfähigkeit einer Mauer falsch berechnet oder eine Übersetzerin einen Text fehlerhaft in eine andere Sprache überträgt, steht grundsätzlich fest, wer für den Schaden haftet. Doch wie verhält es sich, wenn die Fehler durch den Einsatz eines KI-Systems begünstigt wurden? KI wirft neue rechtliche und ethische Fragen auf, die im Folgenden nur angerissen werden können.

Bis jetzt geht die Rechtsordnung davon aus, dass am Anfang einer Handlung die Entscheidung eines Subjekts steht, an welche alles anknüpft, was danach geschieht – einschliesslich der Rechtsfolgen. Diese Vorstellung, wonach ein Schaden seinem Urheber zugerechnet werden kann, wird durch KI infrage gestellt. Das Problem verschärft sich in dem Masse, als KI-Systeme zunehmend eigenständige und unvorhersehbare Entscheidungen treffen.

KI-System an die Leine legen?

Die Vorstellung, für einen Schaden geradestehen zu müssen, den er nicht direkt verursacht hat, ist jedem Hundebesitzer ein Schrecken: Egal, ob der Vierbeiner Löcher im benachbarten Garten gegraben oder nach der Hand der Postbotin geschnappt hat, ist es sein Besitzer, der Entschädigung zahlen oder sonstige Konsequenzen tragen muss. Der Halter kann sich von der Haftung nur befreien, wenn er nachweist, dass er im Umgang mit seinem Tier die nach den Umständen gebotene Sorgfalt hat walten lassen.

Verschiedene Rechtsgelehrte sind der Ansicht, die Tierhalterhaftung könne als Vorbild für eine spezifische Haftung autonomer KI-Systeme dienen: Auch diese sind mitunter unberechenbar und haben das Potenzial, Unheil anzurichten. Kritiker dieser Analogie geben zu bedenken, dass die Tätigkeit der KI, die zum Schaden geführt hat, an die KI delegiert worden ist, der Ursprung der ganzen Handlungskette also bei der Person liege, die ein KI-System eingesetzt hat.

Einem weitergehenden Konzept zufolge könnte KI-Systemen ein eigener rechtlicher Status zuerkannt werden. Die EU-Kommission denkt seit 2017 über die Schaffung einer «elektronischen Person» nach, die es ermöglichen soll, fortgeschrittene Robotersysteme für die von ihnen verursachten Schäden haftbar zu machen; ein ähnlicher Ansatz wäre auch für KI-Systeme vorstellbar, ist aber gemäss der Einschätzung von Fachleuten zumindest gegenwärtig noch klar abzulehnen.

Kreativität aus erster und zweiter Hand

Mathematische Methoden sind nach geltendem Recht von der Patentierbarkeit ausgeschlossen. Weniger klar verhält es sich mit dem Patentschutz von Computerprogrammen. Wenn nämlich maschinelles Lernen einen «technischen Beitrag» im Rahmen einer Erfindung darstellt – etwa, indem neuronale Netze für die Klassifizierung digitaler Bilder und Videos zum Einsatz kommen – kann laut Praxis des Europäischen Patentamts die Patentierbarkeit gegeben sein. Ob also ein KI-System patentiert werden kann, muss von Fall zu Fall geprüft werden.

Auch mit Blick auf einen allfälligen urheberrechtlichen Schutz sind Fragen offen. Computerprogramme gelten zwar als Werke, die urheberrechtlichen Schutz geniessen können. Allerdings ergibt sich ein Problem bei Systemen, die selber lernen und sich dadurch stetig weiterentwickeln. Damit verändert sich auch der Gegenstand des urheberrechtlichen Schutzes – denn ein KI-System ist zu einem späteren Zeitpunkt unter Umständen nicht mehr das gleiche wie zu Beginn.

Eine weitere Schwierigkeit stellt sich bei immateriellen Gütern, die durch KI-Systeme erzeugt werden. Es ist unklar, ob solche KI-generierten Güter überhaupt schutzfähig sind und – wenn ja, wer aus immaterieller Sicht als Schöpfer entsprechender Ergebnisse zu gelten hat: Der Programmierer, der das betreffende System entwickelt hat, die Person, die es einsetzt – oder das System selbst? Während

letztere Möglichkeit bereits aufgrund der mangelnden Rechtsfähigkeit von KI-Systemen ausscheidet, muss in Bezug auf die Zuweisung der betreffenden Schutzrechte an eine natürliche Person differenziert werden:

Unter dem Urheberrecht dürfte die Schutzfähigkeit im Falle von vollständig autonom durch KI generierten Werken bereits daran scheitern, dass kein Mensch geistig zu deren Schöpfung beigetragen hat – im Urheberrecht eine Grundvoraussetzung für die Schutzfähigkeit einer Kreation. Allerdings ist in der Praxis die Abgrenzung solcher Werke von Erzeugnissen, an deren Entstehung KI bloss als Hilfsmittel beteiligt war, mit Schwierigkeiten verbunden.

Aus Sicht des Patentrechts ist hingegen von der grundsätzlichen Schutzfähigkeit KI-generierter Erfindungen auszugehen. Dabei wird jene natürliche Person zum Erfinder im Rechtssinne, welche das Resultat als erste zur Kenntnis nimmt und als Lösung eines technischen Problems begreift.

Fairness lässt sich schwer berechnen

KI wirft nebst rechtlichen auch ethische Probleme auf, die kaum zu lösen sind. Illustrieren lässt sich das am US-amerikanischen KI-System «COMPAS», das anhand von 137 Merkmalen schätzt, ob Kriminelle zu Wiederholungstätern werden. In die Kritik rückte das System, als sich zeigte, dass es bei Afroamerikanern fast doppelt so häufig wie bei Weissen fälschlicherweise ein hohes Rückfallrisiko berechnete. Bei Weissen wiederum wurde ein Rückfall fast doppelt so oft nicht vorausgesagt wie bei Schwarzen. Der Algorithmus hatte dabei keine Angaben über die ethnische Zugehörigkeit der Angeklagten erhalten.

Die Firma, die den Test anbietet, wehrte sich gegen den Vorwurf des Rassismus mit dem Hinweis, sie habe einen Test entwickelt um sicherzustellen, dass die Präzisionsrate für dunkel- und hellhäutige Menschen gleich hoch sei – und das gewährleiste eine faire Software. Mathematiker haben sich in der Folge mit der Frage beschäftigt, ob ein Test zugleich fair und unfair sein kann, und kommen zum Schluss, das sei sehr wohl möglich – nämlich dann, wenn eine Eigenschaft (hier die Rückfälligkeit) in einer Grundgesamtheit je nach Untergruppe unterschiedlich oft auftritt. Da Afroamerikaner öfter wieder straffällig werden als Weisse, sei es nicht möglich, für beide Gruppen gleich präzise Vorhersagen zu treffen und zugleich für beide Ethnien gleich hohe

Fehlerquoten zu erzielen. Gerechtigkeit, so die Fachleute, könne nicht statistisch oder mathematisch hergestellt werden. Vielmehr falle es der Ethik zu, über Vorgehensweisen und Kriterien nachzudenken, um die Fairness von Algorithmen sicherzustellen oder zumindest zu erhöhen.

Ethikrichtlinien auf europäischer Ebene

Im Frühjahr 2019 veröffentlichte eine Gruppe von KI-Experten im Auftrag der Europäischen Kommission Ethikrichtlinien für eine vertrauenswürdige KI. Die Richtlinien sind sowohl für Entwickler von KI als auch für Personen gedacht, die KI einsetzen oder mit deren Ergebnissen arbeiten. Damit soll sichergestellt werden, dass Grundwerte und geltende Vorschriften respektiert werden und der Einsatz der KI einem ethisch erstrebenswerten Ziel verpflichtet bleibt.

Besonders Gewicht legen die Richtlinien darauf, dass die Kontrolle von Entscheidungen und Vorgängen auch bei der Verwendung von KI beim Menschen verbleibt. Technische Systeme sollen robust sein und zuverlässig wiederholbare Ergebnisse liefern. Daten und Privatsphäre sind zu schützen, ausserdem sollen KI-Systeme auf transparente Weise eingesetzt werden, damit die Wege, die zu einer Entscheidung führen, nachvollziehbar bleiben. Schliesslich gilt es, bei der Nutzung von KI-Systemen unfaire Verzerrungen zu vermeiden und auch das gesellschaftliche und ökologische Wohl im Auge zu behalten.

Verzerrungen und Intransparenz fallen ins Gewicht

Aus Sicht einer Mehrheit der befragten Fachleute stellen unausgewogene Datensätze und Intransparenz beim Einsatz von KI erhebliche Risiken dar. Dennoch stehen die Experten der Nutzung von KI grundsätzlich wohlwollend gegenüber und bescheinigen ihr einen positiven Einfluss auf allgemein akzeptierte gesellschaftliche Ziele wie Armutsbekämpfung oder Stärkung der Gesundheit.

Fest steht aus Sicht der Befragten allerdings, dass die Verantwortlichkeit für KI-Entscheidungen beim Menschen liegen muss. Nutzerinnen und Nutzer müssen informiert werden, wenn sie mit KI interagieren, und sie sollten auch befähigt werden zu verstehen, wie die KI zu ihrer Entscheidung gefunden hat.

Empfehlungen: Ein weites Feld gestalten

KI-Systeme kommen in ganz unterschiedlichen Anwendungsfeldern zum Einsatz. Entsprechend weit greifen Empfehlungen zum Umgang mit dieser Technik aus.

KI-Systeme werden für die unterschiedlichsten Aufgaben in ganz verschiedenen Gebieten eingesetzt, die jeweils auch eigene rechtliche Fragen aufwerfen. Die Schaffung eines «KI-Gesetzes», das auf eine einheitliche Regelung des KI-Einsatzes abzielen würde, ist daher nicht angemessen und deswegen abzulehnen.

Spezifische Anwendungen ausleuchten

Wer nutzt ein KI-System wofür genau, auf welchen Daten baut dieses auf, und welche rechtlichen Vorgaben gilt es dabei zu berücksichtigen? Diese Fragen sind zu beantworten, bevor die Chancen und Risiken eines Einsatzes von KI abgewogen werden können. Die Regeln für den Gebrauch von KI müssen daher stets abgestimmt auf den jeweiligen Verwendungszusammenhang festgelegt werden. Auch sind konkrete Probleme und Fehlentwicklungen zu identifizieren und mit geeigneten Vorkehrungen zu beheben; wo die Risikolage unklar ist, sollte die Forschung intensiviert werden, um allfällige Gefahren zu ermitteln.

Hohe Anforderungen an die Behörden

Nimmt der Staat seine hoheitlichen Aufgaben wahr, handelt er gegenüber dem Einzelnen aus einer übergeordneten Stellung. Daher ist es umso wichtiger, dass die Rechtmässigkeit staatlichen Handelns von den Betroffenen überprüft werden kann und KI-Systeme entsprechend transparent eingesetzt werden. Wenn also der Staat KI verwendet, muss er grundsätzlich höheren Ansprüchen genügen als private Anwender.

Transparenz und verdaubare Information

Sind Personen in relevanter Weise von einer KI-Nutzung betroffen, müssen sie auf eine einfache und gut verständliche Weise in Kenntnis gesetzt werden, dass sie es mit einem KI-System zu tun haben. Dabei ist indes zu beachten, dass ein Übermass an Information zu einer gewissen Resignation und damit zur Nachlässigkeit vonseiten der Betroffenen führen kann. Deshalb ist sicherzustellen, dass diese wie auch Organisationen, die sich anwaltschaftlich für Geschädigte einsetzen, auf Nachfrage sämtliche Angaben erhalten, die es gestatten, fehlerhafte Ergebnisse zu erkennen und einzuschätzen.

Gütesiegel statt generelle Marktzulassungen

Die Verwendung von KI-Systemen generell einer Marktzulassung zu unterstellen, wäre angesichts ihrer vielfältigen und unterschiedlich einschneidenden Einsatzmöglichkeiten nicht sachgerecht. Wo – wie bei Medizinalprodukten – bereits jetzt eine Marktzulassung erforderlich ist, umfasst die Prüfung auch die KI. Wenn aber private Initiativen Gütesiegel für KI entwickeln, ist dies zu begrüssen und zu fördern. Organisationen, die sich beispielsweise für den Konsumentenschutz engagieren, sollten befähigt werden, die Qualität solcher Zertifikate besser zu beurteilen.

Nicht bloss auf technische Fragen fokussieren

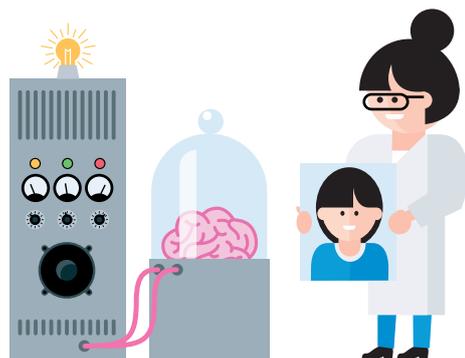
Konzeption und Entwicklung von KI sind bei weitem nicht nur eine Frage der Technik. Insbesondere um die damit verbundenen Risiken einschätzen zu können, ist es unabdingbar, dass Fachleute, welche KI entwickeln, auch über ein ethisches Verständnis sowie grundlegendes juristisches Wissen verfügen und darüber hinaus zur interdisziplinären Zusammenarbeit bereit und fähig sind. Entsprechende Kenntnisse sind auch für all jene unabdingbar, die KI-Systeme einsetzen oder mit ihren Ergebnissen arbeiten.

Mitglieder der Begleitgruppe

- Prof. Dr. Jean Hennebert, Leitungsausschuss TA-SWISS, Département d'informatique de l'Université de Fribourg, Präsident der Begleitgruppe
- Benjamin Bosshard, Eidgenössische Kommission für Kinder- und Jugendfragen
- Sabine Brenner, Geschäftsstelle Digitale Schweiz, Bundesamt für Kommunikation (BAKOM)
- Dr. Christian Busch, Staatssekretariat für Bildung, Forschung und Innovation (SBFI)
- Dr. Christine Clavien, Institut Ethique Histoire Humanités, Université de Genève
- Daniel Egloff, Staatssekretariat für Bildung, Forschung und Innovation SBFI
- Andy Fitze, SwissCognitive – The Global AI Hub.
- Matthias Holenstein, Stiftung Risiko-Dialog
- Dr. Marjory Hunt, Fonds national suisse de la recherche scientifique (FNS)
- Manuel Kugler, Schweizerische Akademie der Technischen Wissenschaften (SATW)
- Thomas Müller, TA-SWISS Leitungsausschuss, Redaktor Schweizer Radio SRF
- Katharina Prelicz-Huber, TA-SWISS Leitungsausschuss (bis 2019), Präsidentin Gewerkschaft VPOD/SSP, Nationalrätin
- Prof. Ursula Sury, Rechtsanwältin und Professorin, Hochschule Luzern (HSLU)
- Dr. Stefan Vannoni, TA-SWISS Leitungsausschuss, cemsuisse

Projektleitung bei TA-SWISS

- Dr. rer. soc. Elisabeth Ehrensperger, Geschäftsführerin
- Dr. Catherine Pugin, Projektleiterin



Impressum

Kopflos klug

Kurzfassung der Studie «Wenn Algorithmen für uns entscheiden:
Chancen und Risiken der künstlichen Intelligenz»

TA-SWISS, Bern 2020

TA 72A/2020

Autorin: Lucienne Rey, TA-SWISS, Bern

Produktion: Fabian Schluep, TA-SWISS, Bern

Gestaltung und Illustrationen: Hannes Saxer, Bern

Druck: Jordi AG – Das Medienhaus, Belp

TA-SWISS – Stiftung für Technologiefolgen-Abschätzung

Neue Technologien bieten oftmals entscheidende Verbesserungen für die Lebensqualität. Zugleich bergen sie mitunter aber auch neuartige Risiken, deren Folgen sich nicht immer von vornherein absehen lassen. Die Stiftung für Technologiefolgen-Abschätzung TA-SWISS untersucht die Chancen und Risiken neuer technologischer Entwicklungen in den Bereichen «Biotechnologie und Medizin», «Informationsgesellschaft» und «Mobilität / Energie / Klima». Ihre Studien richten sich sowohl an die Entscheidungstragenden in Politik und Wirtschaft als auch an die breite Öffentlichkeit. Ausserdem fördert TA-SWISS den Informations- und Meinungsaustausch zwischen Fachleuten aus Wissenschaft, Wirtschaft, Politik und der breiten Bevölkerung durch Mitwirkungsverfahren. Die Studien von TA-SWISS sollen möglichst sachliche, unabhängige und breit abgestützte Informationen zu den Chancen und Risiken neuer Technologien vermitteln. Deshalb werden sie in Absprache mit themenspezifisch zusammengesetzten Expertengruppen erarbeitet. Durch die Fachkompetenz ihrer Mitglieder decken diese Begleitgruppen eine breite Palette von Aspekten der untersuchten Thematik ab.

Die Stiftung TA-SWISS ist ein Kompetenzzentrum der Akademien der Wissenschaften Schweiz.



TA-SWISS
Stiftung für Technologiefolgen-Abschätzung
Brunngasse 36
CH-3011 Bern
info@ta-swiss.ch
www.ta-swiss.ch

mitglied der
 akademien der
wissenschaften schweiz