

DECODING ARTIFICIAL INTELLIGENCE

Warum die Künstliche Intelligenz die menschliche braucht



W.I.R.E.

WEB FOR INTERDISCIPLINARY RESEARCH AND EXPERTISE

—
THINK TANK FOR BUSINESS, SCIENCE AND SOCIETY

NZZ LIBRO

RAPHAEL VON THIESSEN | STEFAN PABST
STEPHAN SIGRIST

DECODING ARTIFICIAL INTELLIGENCE

Warum die Künstliche Intelligenz die menschliche braucht

W.I.R.E. | Think Tank für Wirtschaft, Wissenschaft und Gesellschaft
Im Auftrag der Standortförderung im Amt für Wirtschaft und Arbeit des Kantons Zürich

NZZ LIBRO

DECODING ARTIFICIAL INTELLIGENCE

Der Einsatz von Künstlicher Intelligenz transformiert Wirtschaft, Wissenschaft und Gesellschaft gleichermaßen, denn autonome Computersysteme und selbstlernende Algorithmen werden zu einem festen Bestandteil des Berufs- und Alltagslebens. Im Zentrum des Wandels steht das Versprechen, Produktivität und Lebensqualität durch die Automatisierung von Prozessen und durch die Personalisierung von Produkten und Dienstleistungen nachhaltig und merklich zu steigern. Gleichzeitig stellen sich aus gesellschaftlicher Perspektive richtungsweisende Fragestellungen unter anderem zu den ethischen und regulatorischen Prinzipien für KI-basierte Entscheidungsprozesse, zur Gewährleistung von Transparenz und zur Klärung von Verantwortlichkeiten bei Fehlentscheidungen.

Dabei ist ein Verständnis der technologischen Grundlagen die Voraussetzung für eine realistische Einschätzung der effektiven Potenziale der Künstlichen Intelligenz und für eine gesamtheitliche und differenzierte Betrachtung der Chancen und Herausforderungen: Eine wesentliche Eigenschaft von KI-Systemen ist die automatisierte Erkennung von Mustern aufgrund von statistischen Korrelationen und Strukturen innerhalb von Daten. In Kombination mit spezifischen Lernfähigkeiten zur selbstständigen Optimierung von Ergebnissen lassen sich zentrale Anwendungsbereiche ableiten, in denen Künstliche Intelligenz kurz- und mittelfristig zum Einsatz kommen kann: KI-Systeme ergänzen menschliche Kompetenzen und übernehmen einzelne

Aufgaben, indem sie Empfehlungen abgeben, Handlungsoptionen abwägen und Entscheidungen treffen, aber auch indem sie Vorhersagen und Prognosen optimieren. Darüber hinaus können Computersysteme mit Künstlicher Intelligenz durch die Simulation von Sprachfähigkeit mit Menschen kommunizieren oder diese in der physischen Welt unterstützen, indem sie Roboter oder autonome Vehikel steuern. Letztlich können durch Künstliche Intelligenz von der pharmazeutischen Forschung über die Architektur und Raumplanung bis hin zur Kunst und Kultur neue Lösungsansätze generiert werden, womit die menschliche Kreativität und Innovationskraft ergänzt werden. Die Beurteilung der mittel- bis langfristigen Potenziale der Künstlichen Intelligenz erfordert jedoch auch die Berücksichtigung der Limitationen: So sind Technologien der Künstlichen Intelligenz auf grosse und qualitativ hochwertige Datenmengen sowie hohe Rechenleistungen angewiesen. Weitere Einschränkungen betreffen die Unfähigkeit von KI-Systemen, irrelevante Korrelationen von kausalen Zusammenhängen zu unterscheiden, ebenso wie die fehlende Nachvollziehbarkeit von KI-basierten Ergebnissen. Die Künstliche Intelligenz ist demzufolge nicht ein übermächtiges Allzweckmittel, sondern sie entfaltet ihre Potenziale nur in eng begrenzten Bereichen – und im Zusammenspiel mit der menschlichen Intelligenz.

THESEN

Basierend auf den heutigen sowie den potenziellen Anwendungsfeldern der Künstlichen Intelligenz, gilt es die Zukunft der lernenden Maschinen in einer grösseren Perspektive weiterzudenken:

TOASTERS, NOT TERMINATORS

Intelligente Maschinen werden zur Normalität, ihre Kompetenzen sind jedoch limitiert. KI-Anwendungen werden immer mehr Bereiche des Berufs- und Alltagslebens prägen, sie sind aber keine Kopie menschlicher Intelligenz. Es handelt sich vielmehr um Unterstützungs- und Cockpitsysteme und damit um intelligente Toaster – und nicht Terminator-Roboter.

MACHINE TEACHING, NOT LEARNING

Künstliche Intelligenz erfordert Interaktion mit Menschen und der physischen Umgebung. KI-Systeme lösen Herausforderungen basierend auf einem existierenden Wissensschatz und der Wechselwirkung mit Menschen, die durch Machine Teaching den intelligenten Maschinen Lösungswege beibringen. Die Potenziale entfalten sich also nur, wenn KI-Systeme im Zusammenspiel mit Menschen gedacht, entwickelt und umgesetzt werden.

HUMAN, NOT ALGORITHMIC CONTROL

Die nächste Stufe der Automatisierung entlastet den Menschen, mindert aber seine Entscheidungskompetenz. Der Einsatz von KI-Systemen hilft, in komple-

xen und unübersichtlichen Situationen bessere Entscheidungen zu treffen. In der Tendenz trauen sich Menschen gleichzeitig jedoch immer weniger, eigenständig und intuitiv Urteile zu fällen, was zu einem Autonomie- und Kontrollverlust führen kann.

BOOSTING INDIVIDUALS, SHRINKING INDEPENDENCE

Kompetenzen und Kapazitäten von kleinen Organisationen und Individuen werden gestärkt, indem leistungsfähige KI-Systeme losgelöst von komplexen und kostenintensiven Infrastrukturen menschliche Expertise unterstützen. Gleichzeitig wächst dadurch die Abhängigkeit von KI-Technologien, die massgeblich zur Herstellung und Erbringung von hochwertigen Produkten und Dienstleistungen beitragen.

ACCEPTING BLACK BOXES

Neuronale Netzwerke ermöglichen die Entdeckung neuer Lösungsansätze in Wirtschaft und Wissenschaft, sie erfordern aber auch den Umgang mit Intransparenz, da die Mechanismen hinter den Ergebnissen oftmals nicht nachvollziehbar sind. Im Kontext der Komplexität dieser Systeme stellen sich grundsätzliche Fragen im Hinblick auf regulatorische Massnahmen, demokratische Prozesse und die gesellschaftliche Akzeptanz von KI-basierten Fehleinschätzungen.

TOWARDS AN ECONOMY OF TRUST

Simulationen der Realität schaffen neue Erlebnisse und ermöglichen das Eintau-

chen in immersive Welten. Sogenannte Deep Fakes liefern aber auch die Basis für Verunsicherung aufgrund von gefälschten und manipulierten Medieninhalten im virtuellen Raum. Der Aufbau von Vertrauen wird dadurch zur Grundlage für einen wünschenswerten Einsatz von KI-Technologien in Wirtschaft und Gesellschaft.

EQUALITY IS DEAD. LONG LIVE EQUALITY

Intelligente Algorithmen können Fairness steigern, indem sie Entscheidungsprozesse objektivieren und bestehende Verzerrungen aufdecken. Gleichzeitig kann Künstliche Intelligenz aber ebenso Diskriminierung verschärfen, wenn Zusammenhänge, Positionen oder Einstellungen von Menschen extremer als in der Realität dargestellt werden oder wenn historische Verzerrungen repliziert werden.

FROM BIG DATA TO SMALL DATA

Die nächste Generation Künstlicher Intelligenz wird sich am menschlichen Gehirn orientieren und soll nicht mehr standardisierte, sondern vielfältige Lösungen liefern. Der bisherige Schwerpunkt der KI-Forschung lag dabei auf der Auswertung von Big Data. Durch einen Übergang zu Small Data könnten KI-Systeme künftig auch aus kleineren Datenmengen automatisiert Entscheidungsgrundlagen und Rückschlüsse ableiten.

HANDLUNGSFELDER

Für die Gestaltung einer wünschbaren Zukunft, in der Künstliche Intelligenz echte Mehrwerte generiert, sind Politik, Wirtschaft und Wissenschaft mit vielfältigen Fragestellungen konfrontiert, um nachhaltige Rahmenbedingungen zu schaffen:

TECHNISCHE INFRASTRUKTUR AUFBAUEN

Um durch Künstliche Intelligenz möglichst vielfältige Mehrwerte für Wirtschaft und Gesellschaft zu generieren, sollte der Zugang zu einer grundlegenden KI-Infrastruktur mit hohen Rechenleistungen und Speicherkapazitäten nicht nur Technologiekonzernen, sondern auch verschiedenen Institutionen wie Forschungseinrichtungen, Start-ups, Spin-offs oder KMU offenstehen.

DATENBANKEN UND -STRATEGIEN ENTWICKELN

Der Zugang zu qualitativ hochstehenden Daten ermöglicht Spitzenforschung und die Entwicklung nachhaltiger Praxisanwendungen. Voraussetzung ist der Aufbau von stabilen und vertrauenswürdigen Datenbanken, die unter Berücksichtigung der informationellen Selbstbestimmung den Datenaustausch mit anderen Akteuren fördern.

EIN NEUES NARRATIV FÜR KÜNSTLICHE INTELLIGENZ ENTWICKELN

Das aktuelle Verständnis von Künstlicher Intelligenz fokussiert stark auf die Imitation menschlicher Fähigkeiten. Eine Neupositionierung von Künstlicher Intelligenz als Grundlage für komplementäre und unterstützende Systeme, die ergänzend zu menschlichen Kompetenzen funktionieren, beugt Ängsten und Verunsicherung vor.

BEVÖLKERUNG SENSIBILISIEREN UND VERTRAUEN AUFBAUEN

Eine effektive Sensibilisierung breiter Bevölkerungsschichten setzt ein grundlegendes Verständnis für die technologischen Funktionsweisen von Künstlicher Intelligenz voraus. Der Aufbau von Vertrauen basiert dabei auf der Schaffung von Transparenz beim Einsatz von KI-Systemen und bildet die Grundlage für gesellschaftliche Akzeptanz.

BILDUNG UND KOMPETENZEN AUF DAS ZEITALTER DER KÜNSTLICHEN INTELLIGENZ AUSRICHTEN

Der Einsatz von Künstlicher Intelligenz verändert Berufsbilder kontinuierlich. Die künftig erforderlichen Kompetenzen werden massgeblich durch Agilität und Anpassungsfähigkeit bestimmt sein, weshalb Weiterbildungen in sämtlichen Lebensabschnitten an Bedeutung gewinnen und die Anpassung der Wissens- und Kompetenzvermittlung auf allen Bildungsebenen erforderlich ist.

POSITIONIERUNG IM INTERNATIONALEN WETTBEWERB STÄRKEN

Ein Standort kann sich im Kontext der Künstlichen Intelligenz im globalen Wettbewerb differenzieren, indem er auf ein Teilgebiet fokussiert und dadurch die akademischen und ökonomischen Kräfte bündelt. Im europäischen Kontext kann der Weg darin liegen, sich neben der technologischen Exzellenz auf den gesellschaftlichen Nutzen zu konzentrieren und beispielsweise einen Small-Data-Ansatz voranzutreiben, um sich im Wettbewerb mit bevölkerungsreichen Staaten wie den USA oder China zu differenzieren.

GESELLSCHAFTLICHE UND ETHISCHE GRUNDLAGEN KLÄREN

Durch den Einsatz von KI-Systemen wächst die Notwendigkeit für die Bestimmung von ethischen Grundlagen. Im Zentrum stehen Transparenz, Gleichberechtigung, Sicherheit und Selbstbestimmung. Neben der Erarbeitung von Richtlinien und Standards braucht es einen gesellschaftlichen Konsens darüber, in welchen Bereichen künstlich intelligente Systeme wünschenswert sind.

ÖKOSYSTEME AUFBAUEN

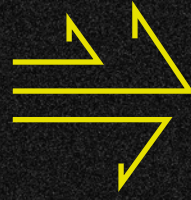
Aufgrund der vielfältigen Chancen und komplexen Herausforderungen einer KI-basierten Ökonomie und Gesellschaft gilt es die übergreifende Zusammenarbeit zwischen Politik, Wirtschaft und Wissenschaft zu fördern, um Potenziale der

Künstlichen Intelligenz zu stärken und gleichzeitig die Risiken zu minimieren. Die geografische Nähe zwischen den Akteuren und die Flexibilität der transdisziplinären Netzwerke sind zentrale Erfolgsfaktoren eines effizienten und zielgerichteten Aufbaus von KI-Ökosystemen.

NEUE AUFGABEN FÜR DIE ÖFFENTLICHE HAND DEFINIEREN

Um punktuellen und einseitigen Reaktionen vorzubeugen, kann die Politik kritische Fragen im Kontext der Künstlichen Intelligenz frühzeitig in den öffentlichen Dialog bringen. Aufgrund der Komplexität und Vielseitigkeit von KI-basierten Lösungen können darüber hinaus gestaltende regulatorische Gefässe geschaffen werden, in denen unter strenger Beaufsichtigung und Berücksichtigung lokaler Gegebenheiten kontrolliert experimentiert wird, um im Anschluss praxistaugliche regulatorische Rahmenbedingungen zu schaffen.

10



1

Einleitung

ZUM AUFSTIEG INTELLIGENTER MASCHINEN

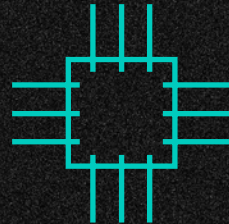


4

Konsequenzen

WAS KI FÜR DIE WELT VON MORGEN BEDEUTET

18



2

Technologische Grundlagen

WIE MASCHINEN MUSTER ERKENNEN UND LERNEN



5

Thesen

ZUR ZUKUNFT INTELLIGENTER MASCHINEN

32



3

Anwendungsbereiche

WAS MASCHINEN KÖNNEN



6

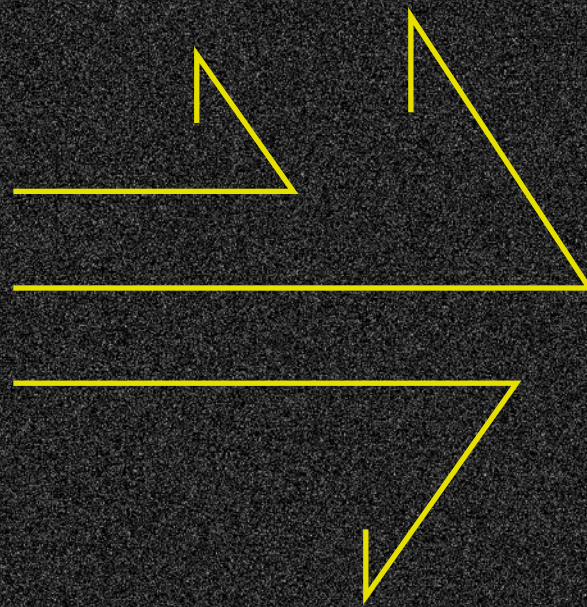
Handlungsfelder

**AUFGABEN ZUR GESTALTUNG EINER
WÜNSCHBAREN ZUKUNFT**

48

56

74



1

Einleitung

ZUM AUFSTIEG INTELLIGENTER MASCHINEN

Die Nachbildung von intelligentem Verhalten ist ein Menschheitstraum. Seit Jahrhunderten beschäftigt sich die Menschheit mit der Vision, eines Tages einen künstlichen Menschen zu schaffen. Mit den rasanten Fortschritten der Computerwissenschaften scheint dieser Traum nun Realität zu werden. Was bis vor Kurzem nur als literarische Erzählung oder cineastische Inszenierung erlebbar war, wird nun Teil des Alltags: Lernende und selbstständig entscheidende Computersysteme, die – so das Versprechen von Forschung und Unternehmen – immer mehr Aufgaben von Menschen unterstützen oder sie ihnen gänzlich abnehmen. Dies würde für die Welt bereits heute eine tief greifende Zäsur darstellen. Denn wenn der Mensch und sein Kern – nämlich die Fähigkeit zu denken – durch Maschinen reproduziert werden können, führt dies zu grossen Hoffnungen, aber auch Ängsten. Denn bereits heute verändern lernende Algorithmen die Art und Weise, wie Alltagsentscheidungen getroffen werden und prägen die Geschäftsmodelle und Kundeninteraktionen von Banken über Mobilitätsanbieter bis zum Gesundheitssystem. Dieser Einfluss wird weiter wachsen und erfordert heute eine offene, aber auch kritische Auseinandersetzung mit den Chancen und Herausforderungen der Künstlichen Intelligenz – in der Politik, der Wirtschaft und der Öffentlichkeit. Eine differenzierte Perspektive und Strategie, die lernende Maschinen jenseits der aktuellen Hypes dort einsetzen, wo sie Nutzen stiften und kritische Entwicklungen berücksichtigt, legen die Grundlage für die Gestaltung einer prosperierenden und nachhaltigen Gesellschaft mit hoher Lebensqualität.



Die vermeldeten Durchbrüche auf dem Gebiet der Künstlichen Intelligenz (KI) sind nicht nur medial äusserst präsent, sondern dringen auch immer weiter in den Alltag und das Arbeitsleben breiter Bevölkerungsschichten hinein. Nach Jahren grossartiger Visionen erreichen Anwendungen der Künstlichen Intelligenz zumindest in eng begrenzten Aufgabenbereichen das kognitive Leistungsniveau von Menschen oder übertreffen dieses sogar: Medizinische Diagnosen durch Bilderkennungsprogramme, personalisierte Produktempfehlungen oder automatisierte Übersetzungen beschreiben nicht mehr länger ambitionierte Forschungsziele, sondern bilden den Kern aktueller Dienstleistungen und Geschäftsmodelle. Angesichts der rasanten technologischen Entwicklungen und hohen Dynamik der Forschung im Umfeld intelligenter Maschinen ist es nicht weiter erstaunlich, dass der gesellschaftliche Diskurs zwischen Euphorie und Untergangsszenarien schwankt. Auf der einen Seite erhoffen sich Optimisten und Optimistinnen eine drastische Verbesserung der Lebensqualität durch mehr Effizienz und Personalisierung in sämtlichen Lebensbereichen. Auf der anderen Seite des Spektrums verkünden Kritiker und Kritikerinnen die baldige Machtübernahme einer Superintelligenz oder die Dystopie eines KI-basierten Überwachungsstaats. Die Debatten

rund um das Thema Künstliche Intelligenz werden in der Konsequenz immer unüberschaubarer und die Orientierung wird sowohl für Experten und Expertinnen als auch für Bürger und Bürgerinnen zunehmend schwieriger.

Künstliche Intelligenz verstehen

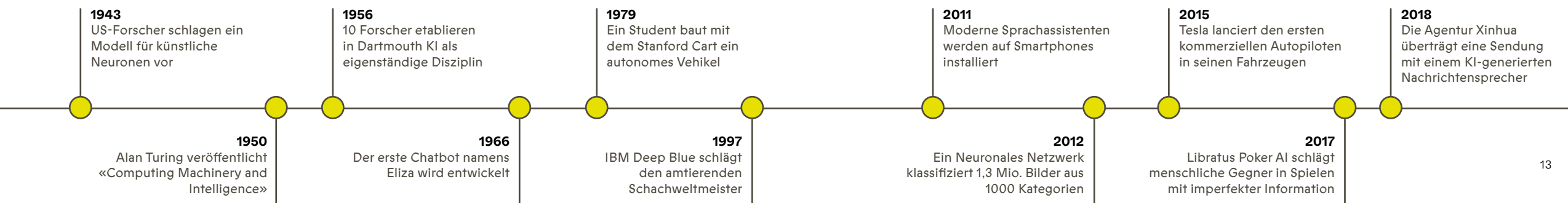
Dabei könnte alleine die begriffliche Unterscheidung zwischen Starker und Schwacher Künstlicher Intelligenz zu mehr Klarheit und Schärfe in der gesellschaftlichen Debatte beitragen. Der US-amerikanische Sprachwissenschaftler John Searle hat diese Trennung bereits vor 50 Jahren vorgeschlagen, um den schon damals umstrittenen Begriff der Künstlichen Intelligenz zu strukturieren.¹ Das Forschungsfeld der Starken Künstlichen Intelligenz hat die Zielsetzung, eine allgemeine und menschenähnliche Intelligenz zu schaffen, die sich auf eine theoretisch unbegrenzte Anzahl von Problemstellungen anwenden lässt. Dieser Ansatz wirft sogleich grundsätzlichere Fragen nach dem Bewusstsein oder der Selbstbestimmung von Maschinen auf. Das Forschungsfeld der Schwachen Künstlichen Intelligenz fokussiert hingegen auf stark eingeschränkte Aufgabenbereiche mit hohem Spezialisierungsgrad und zielt darauf ab, einzelne Anwendungsprobleme wie Routenplanung, Übersetzung oder Bilderkennung automatisiert zu lösen. Damit

lässt sich die Schwache Künstliche Intelligenz klar von einer umfassenden menschlichen Intelligenz abgrenzen. Die meisten Ängste und Dystopien basieren auf der Annahme, dass in Zukunft eine Starke Künstliche Intelligenz geschaffen werden könnte. Dabei ist eine realistische und differenzierte Perspektive gefragt: Aktuell stammen die verkündeten Durchbrüche fast ausschliesslich aus dem Bereich der Schwachen Künstlichen Intelligenz. Eine allgemeine, menschenähnliche Intelligenz zu schaffen, stellt technologisch eine grosse Herausforderung dar und die Realisierung eines solchen Vorhabens ist zumindest aus heutiger Perspektive nicht absehbar. Um die gesellschaftlichen Technikfolgen zu antizipieren, ist es erforderlich, frühzeitig die Weichen zu stellen und sich mit den Implikationen einer möglichen Superintelligenz auseinanderzusetzen. Die vorliegende Publikation betrachtet jedoch bewusst die kurz- bis mittelfristigen Entwicklungen auf dem Gebiet der Künstlichen Intelligenz, um die Dringlichkeit der aktuellen Chancen und Herausforderungen zu adressieren.

Wettbewerb der KI-Supermächte

Insbesondere in der Politik ist das Thema Künstliche Intelligenz angekommen. Aufgrund der vielseitigen Potenziale konkurrieren sich Forschungs- und Entwicklungsstandorte in einem globalen Wettbewerb.

Viele Länder, Regionen und Städte erlassen KI-Strategien und ringen um Forschungsgelder, Firmen und Talente. Allen voran die beiden KI-Supermächte China und USA nutzen dabei die Technologie als marktwirtschaftlichen Treiber, aber auch als Mittel zum staatlichen Machterhalt. Das Ringen um die Vormachtstellung erinnert dabei an die Reise zum Mond und damit an eine Demonstration der Überlegenheit des eigenen Systems: Während China mit einem langfristigen Plan und hohen staatlichen Investitionen operiert, setzen die USA auf die Innovationskraft ihrer Technologiekonzerne und Start-ups und die damit verbundene Venture-Capital-Kultur. Zu kurz kommen dabei bislang Debatten zu den gesellschaftlichen Herausforderungen, die durch den rasanten technologischen Wandel entstehen. Die Rahmenbedingungen und Regulierungen flexibel den fortschreitenden Innovationen im Bereich Künstliche Intelligenz anzupassen, wird zum Schlüsselfaktor für die Gestaltung einer erfolgreichen und positiven Zukunft. Um aus der aktuellen Situation heraus die künftigen Fragestellungen anzugehen, lohnt sich ein kurzer Blick auf die Entstehungsgeschichte der Künstlichen Intelligenz als eigenständiges Forschungsgebiet.



**Turings Vermächtnis**

Bereits gegen Ende des Zweiten Weltkriegs hatten die beiden US-Forscher Warren McCulloch and Walter Pitts die Idee, Konzepte der Logik, Neurologie und Informatik miteinander zu verbinden und intelligentes Verhalten basierend auf künstlichen Neuronen zu simulieren.² Einige Jahre später hat Alan Turing mit seinem Artikel «Computing Machinery and Intelligence» viele Ideen eingeführt, die bis zum heutigen Tag ihre Relevanz behalten. Im Zentrum stand dabei der Gedanke, «nicht den Geist eines Erwachsenen nachzubilden, sondern denjenigen eines Kindes zu imitieren».³ Im Sommer 1956 haben zehn führende Wissenschaftler und Experten aus der Praxis im Rahmen eines Forschungsprojekts zwei Monate lang intensiv im US-Bundesstaat New Hampshire zusammengearbeitet und die Grundsteine für die Künstliche Intelligenz als eigenständige akademische Disziplin gelegt. Die Forscher waren sich darüber einig, dass Ideen, Perspektiven und Techniken aus unterschiedlichen Wissenschaften wie der Philosophie, Mathematik, Informatik, Ökonomie, Neurologie, Psychologie und Linguistik miteinander kombiniert werden müssen, um das Fernziel der Konstruktion intelligenter Maschinen zu erreichen.

Frühe Erfolge bei der Verwendung von Computerprogrammen zur Lösung von einfachen algebraischen oder geometrischen Aufgaben resultierten in den 1950er- und 1960er-Jahren in grossem Enthusiasmus und gewagten Prognosen. Der amerikanische Sozialwissenschaftler Herbert Simon war bereits 1957 davon überzeugt, dass in-

Künstliche Intelligenz – eine Frage der Definition

Die Begriffsdefinition von Künstlicher Intelligenz ist umstritten. Für viele Definitionen steht die Automatisierung von intelligentem Verhalten im Vordergrund, wobei bereits für Intelligenz selbst keine eindeutig abgrenzbare Definition besteht. In der vorliegenden Publikation wird Künstliche Intelligenz als die Fähigkeit von Computersystemen aus Soft- und Hardwarekomponenten verstanden, Aufgaben erfolgreich zu bearbeiten, wobei der Weg zur Lösung der Aufgaben weder statisch noch definitiv vorgegeben ist.

nerhalb von zehn Jahren ein Computer den Schachweltmeister schlagen würde, was jedoch erst kurz vor der Jahrtausendwende gelang.⁴ Die Zuversicht wurde gebremst, als die Systeme an komplexeren Aufgaben wie der Sprachverarbeitung scheiterten. In der Folge kam es zu mehreren sogenannten KI-Wintern, in denen die öffentlichen und privaten Investitionen und damit auch die Forschungsbemühungen auf dem Gebiet der Künstlichen Intelligenz deutlich zurückgingen. Erst nach der Jahrtausendwende begann eine neue Dynamik die Forschung und Entwicklung zu prägen.

Bytes, Daten und selbsterfüllende Prophezeiungen

Doch wieso gewann das Thema Künstliche Intelligenz innerhalb von wenigen Jahren derart an Relevanz? Verantwortlich für Durchbrüche wie die automatisierte Bilderkennung oder die Installation von Autopi-

«Die Faszination für KI ist mit dem Begriff der Intelligenz verbunden, weil dieser Assoziationen weckt, die weit über technische Instrumente hinausgehen.»

Eva Weber-Guskar, Philosophieprofessorin und Fellow der Digital Society Initiative der Universität Zürich

loten in selbstfahrenden Fahrzeugen sind vor allem zwei Haupttreiber, die den rasanten technologischen Fortschritt begünstigten: mehr Rechenleistung und verbesserte Datenverfügbarkeit. Die beiden Treiber ebneten den Weg für die Entwicklung von vielfältigen Anwendungen in den Bereichen Gesundheit, Mobilität, Finanzdienstleistungen, Bildung oder Einzelhandel und führten dazu, dass sich immer mehr Institutionen und Unternehmen dem Thema der Künstlichen Intelligenz annehmen.

Der erste Treiber, die erhöhte Rechenleistung, beruht auf einer der wenigen Technologieprognosen, die sich über Jahrzehnte hinweg bewahrheiteten. Die Einhaltung von Moore's Law ermöglichte es Computersystemen, auch im Kontext der Künstlichen Intelligenz schneller und kostengünstiger immer komplexere Aufgaben zu bearbeiten. Grund für den konstanten Anstieg der verfügbaren Rechenleistung war, dass sich die Anzahl der Transistoren auf Mikrochips wie von Gordon Moore prognostiziert seit den 1960er-Jahren alle zwei Jahre verdoppelte.⁵ Doch erst in Kombination mit dem zweiten Treiber, der verbesser-

ten Datenverfügbarkeit, konnte sich das Potenzial der Künstlichen Intelligenz entfalten. Auslöser für die Generierung immer grösserer Datenmengen war die Durchdringung von Wirtschaft, Wissenschaft und Gesellschaft durch digitale Technologien. Im Zentrum der Digitalisierung stehen neben der Erhebung von Daten auch deren Verarbeitung, Speicherung und Übermittlung. Mit dem Aufkommen des Internets sind die verfügbaren Datenmengen noch einmal exponentiell angestiegen. Die Quantifizierung von sämtlichen Lebens- und Arbeitsbereichen bildet dabei das Fundament für die Entwicklung von aktuellen und künftigen KI-basierten Anwendungen.

L'Homme Machine

Die aktuelle Faszination und der Aufruhr rund um Künstliche Intelligenz lässt sich allerdings nicht ausschliesslich technologisch begründen. Die Vorstellung, dass sich die Mechanismen menschlichen Denkens mechanisch nachbilden und automatisieren lassen, um eine dem Menschen ebenbürtige Maschine zu bauen, ist Teil eines sehr alten Traums der Menschheit. Dieser hatte im Zug der Aufklärung und des wachsenden Ver-



ständnisses der menschlichen Biologie neuen Auftrieb gewonnen. Julien Offray de La Mettrie veröffentlichte 1748 seine Idee eines «Homme Machine», die den menschlichen Organismus als Maschine interpretierte. In seiner Zeit galt dies als Angriff auf das, was den Menschen ausmacht: Vernunft, Geist, Seele, Mortalität, aber auch Glück und Freiheit, die in dieser Vorstellung zu Determinanten der Biologie werden. Fragestellungen, die allerdings auch im 21. Jahrhundert nicht an Brisanz verloren haben. Denn wenn Computerprogramme wie kurz vor der Jahrtausendwende amtierende Schachweltmeister schlagen oder wie im Jahr 2011 weltweit führende Jeopardy-Spieler besiegen, dann legen hoffnungsvolle Beobachter und Beobachterinnen diese Durchbrüche als vermeintliche Bestätigung für das Anbrechen einer neuen Evolutionsstufe aus. Andererseits schwingt bei Betrachtern und Betrachterinnen mit einer pessimistischeren Perspektive auch immer die Urangst mit, von der eigenen Schöpfung übertrumpft zu werden. Gerade diese Bilder sind es, die auch kulturgeschichtlich mit künstlichen Menschen verknüpft wurden. Von Olimpia in E. T. A. Hoffmanns «Der Sandmann» bis Mary Shelleys «Frankenstein» oder bis zum durch Arnold Schwarzenegger verkörperten Terminator: Das Spiel mit der Nachahmung des Menschen bringt Gefahren durch Unterjochung oder den Verlust des eigenen Verstands.

Der Aufstieg der Künstlichen Intelligenz als Teil der Wirtschaft, Gesellschaft und Politik von morgen muss und kann vor diesem Hintergrund nicht allein aus einer techni-

schen und marktwirtschaftlichen Perspektive betrachtet werden. Im Gegenteil, die rasanten technischen Fortschritte erfordern eine Kopplung mit den gesellschaftlichen Rahmenbedingung und eine Sensibilisierung mit Blick auf Hoffnungen und Befürchtungen der Menschen.

Beyond the Hype

Die Grundlage für diese Prozesse ist der Aufbau eines differenzierten Verständnisses für die echten Potenziale und Risiken Künstlicher Intelligenz – abseits der wahrscheinlich naiven und einseitigen Vorstellung, dass superintelligente Roboter im Begriff sind, die Macht zu übernehmen. Das «decodieren» des aktuellen Hypes bildet eine erste zentrale Grundlage für die Entwicklung nachhaltiger Strategien. Denn auch wenn viele der heutigen Bilder und Vorstellungen zu weit gehen und nicht dem entsprechen, was Maschinen künftig tatsächlich können, ist Künstliche Intelligenz mit hoher Wahrscheinlichkeit nicht einfach ein vorübergehender Trend. Im Gegenteil: Die digitale Infrastruktur und die datenbasierte Wirtschaft liefern die nötigen Rahmenbedingungen für eine kontinuierliche Weiterentwicklung von KI-Technologien. Weil aktuelle und kommende Innovationen den Alltag immer weiter durchdringen und gleichzeitig auch Berufsbilder verändern, ist es aus Sicht von Politik, Wirtschaft, Wissenschaft und Gesellschaft äusserst relevant, sich der Thematik frühzeitig anzunehmen. Um die Zukunft zu gestalten, gilt es sich mit einigen wichtigen Fragen auseinanderzusetzen. Auf welchen Prinzipien beruhen Technologien der Künstlichen Intelligenz? Was sind die

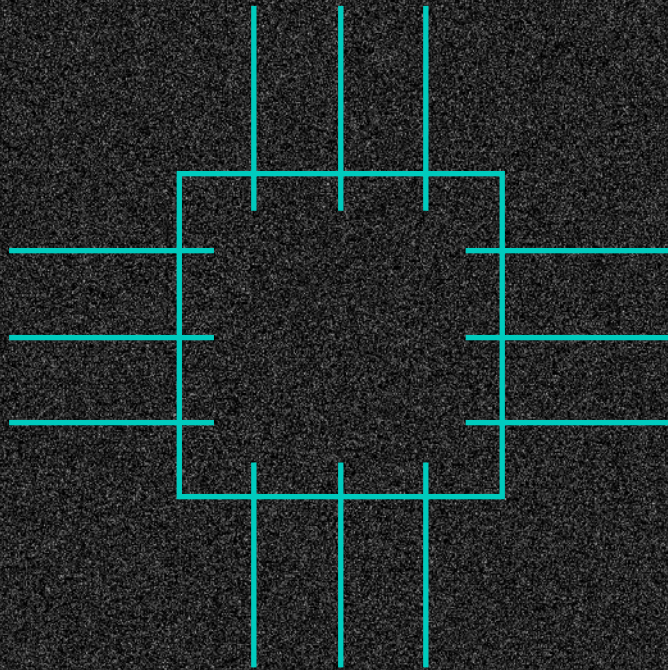
wichtigsten Anwendungsmöglichkeiten? Wo liegen die grössten Chancen? Welche Herausforderungen stellen sich aus gesellschaftlicher Perspektive? Mit welchen Massnahmen können die verschiedenen Akteure und Akteurinnen aus Politik, Wirtschaft und Wissenschaft die Rahmenbedingungen für Technologien der Künstlichen Intelligenz positiv beeinflussen? Mit diesen Fragen beschäftigt sich die vorliegende Publikation mit einem differenzierten Blickwinkel.

Eine wünschbare Zukunft gestalten

Die Grundlage für die Auseinandersetzung mit und die Gestaltung einer Zukunft, in der Künstliche Intelligenz Teil des Alltags wird, bildet das Verständnis der Funktionsweisen lernender Maschinen. Das nachfolgende Kapitel erläutert die zentralen Eckpfeiler Künstlicher Intelligenz: von der Beschreibung automatisierter Mustererkennung bis zu den Mechanismen der digitalen Lernfähigkeit. In einem nächsten Schritt wird ein Modell entwickelt, das die verschiedenen Anwendungsbereiche von Künstlicher Intelligenz zusammenfasst und diese anhand von praktischen Beispielen aus unterschiedlichen Sektoren illustriert. Das vierte Kapitel widmet sich den kurz- bis mittelfristigen Chancen und Herausforderungen, die durch den Einsatz von Künstlicher Intelligenz entstehen. Darauf basierend werden Thesen abgeleitet, die die zukünftige Rolle der Technologie pointiert zum Ausdruck bringen, um in einem letzten Schritt konkrete Handlungsfelder für Politik, Wirtschaft und Wissenschaft aufzuzeigen.

Die Inhalte dieser Studie wurden im Auftrag der Standortförderung im Amt für Wirtschaft und Arbeit des Kantons Zürich erarbeitet. Gleichzeitig wurden ausgewählte Experten und Expertinnen beigezogen, die mit ihren Inhalten dazu beigetragen haben, die Erkenntnisse rund um Künstliche Intelligenz zu erweitern und zu schärfen. Insbesondere Prof. Thomas Hofmann und Prof. Dirk Helbing von der ETH Zürich, Prof. Jana Koehler vom Deutschen Forschungszentrum für Künstliche Intelligenz, Prof. Eva Weber-Guskar von der Digital Society Initiative der Universität Zürich, Dr. Gabriel Krummenacher von der Zühlke Engineering AG und Pascal Kaufmann von der Stiftung Mindfire sei an dieser Stelle für die konstruktive Zusammenarbeit ausdrücklich gedankt.

Künstliche Intelligenz verspricht Mehrwerte für Individuen, Unternehmen, Umwelt und die Gesellschaft als Ganzes. Ein zu starker Fokus auf technologische Potenziale und betriebswirtschaftliche Rationalität lassen dabei schnell vergessen, dass der Mensch im Zentrum dieser Entwicklung steht. Ein gesellschaftlicher Diskurs sollte also auch unbedingt der Wünschbarkeit künftiger Szenarien genügend Raum beimessen. Deshalb dient diese Studie als Grundlage und Inspiration für ergebnisoffene und kritische Debatten zum Thema Künstliche Intelligenz – und als Basis für das Erarbeiten von langfristigen und verbindlichen Strategien und Handlungsfeldern im Umgang mit intelligenten – und weniger intelligenten – künstlichen Systemen, die Teil der Welt von heute und morgen sind.

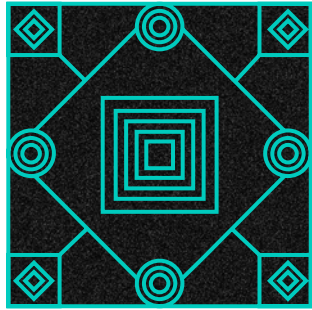
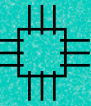


2

Technologische Grundlagen

WIE MASCHINEN MUSTER ERKENNEN UND LERNEN

Hinter den praktischen Anwendungen von Künstlicher Intelligenz stehen Basistechnologien wie Machine Learning, Computer Vision oder Natural Language Processing. Sie legen das Fundament für das, was intelligente Maschinen ausmacht: die Fähigkeit, Muster in grossen Datenmengen zu erkennen – und, noch zentraler, Prozesse selbstständig zu erlernen und zu optimieren. Es existieren aber auch technische und statistische Limitationen, deren Berücksichtigung für eine realistische Beurteilung der effektiven Leistungsfähigkeit von Künstlicher Intelligenz notwendig ist. Eine solche Einschätzung ist elementar, um mit Blick auf die Entwicklung und Lancierung neuer KI-Systeme langfristig erfolgreiche strategische Entscheidungen zu treffen.



WIE MASCHINEN MUSTER ERKENNEN

Die grundlegende Fähigkeit zur Klassifizierung von Informationen ist der Grundstein von Begriffsbildung, Abstraktion sowie induktivem Denken und damit von menschlicher Intelligenz. Sie ermöglicht es, Ordnung in die oftmals chaotische Vielfalt der Sinneswahrnehmungen zu bringen. Diese Anforderung trifft auch auf künstlich intelligente Systeme zu: Eine grundlegende Fähigkeit, die Systeme der Künstlichen Intelligenz auszeichnet, ist das Identifizieren von Strukturen durch Regelmässigkeiten und Muster innerhalb von Datensätzen.¹

Um diese Fähigkeit in einen Nutzen zu übersetzen, gilt es zunächst Informationen aus der realen Welt in eine digitale Form umzuwandeln, beispielsweise indem visuelle oder auditive Signale digitalisiert werden, damit Computersysteme die Daten verarbeiten können. Bei der Mustererkennung wird nach statistischen Korrelationen, Ähnlichkeiten, Regelmässigkeiten oder Wiederholungen gesucht. In zwei For-

schungsfeldern, die Künstliche Intelligenz auszeichnen, wurden in den letzten Jahren massgebliche Fortschritte erzielt. Mit Blick auf die künftigen Potenziale intelligenter Maschinen gilt es, diese genauer zu beleuchten.

Computer Vision: Visuelle Mustererkennung

Das menschliche Gehirn verarbeitet pro Sekunde elf Millionen Sinneseindrücke, wovon lediglich 50 bewusst wahrgenommen werden.² Dabei ist der menschliche Sehsinn für die Orientierung in der physischen Welt von zentraler Bedeutung. Das Forschungsgebiet Computer Vision (zu Deutsch: maschinelles Sehen) versucht, diese Fähigkeiten in digitalen Systemen zu simulieren. Zentrale Verfahren umfassen die Kontrastanalyse zur Erkennung und Klassifikation von geometrischen Objekten oder die Analyse des optischen Flusses, um durch die Zuordnung von Punkten zu Objekten Bewegungsinformationen auf Bildmaterial zu interpretieren. Darüber hinaus werden Farbklassifikatoren eingesetzt, um beispielsweise menschliche Hautfarben zu identifizieren. Mithilfe dieser Verfahren wird versucht, die Mehrdimensionalität des menschlichen Sehens zu imitieren. Allerdings ist die Simulation der visuellen Wahrnehmung zur Bild- und Videoerkennung aufgrund dieser Vielschichtigkeit ausgesprochen komplex. Die digitale Bilderkennung beruht auf der Identifikation von Pixelpunkten. Somit versucht das KI-System anhand von visuellen Eigenschaften wie Konturen oder Farben Objekte zu erkennen und durch den Abgleich mit bestehenden

Daten zu klassifizieren. Die Bild- und Videoerkennungsprogramme können in eng begrenzten Kategorien zuverlässig Klassifikationen vornehmen, wie beispielsweise in der Logistik zur Erkennung von Handschriften bei der Auswertung von händisch verfassten Dokumenten, bei selbstfahrenden Fahrzeugen zur Warnung vor Fussgängern oder bei Sicherheitsfirmen zur Identifikation von Gesichtern. Eine der relevantesten Limitationen besteht jedoch darin, dass die selbstlernenden Algorithmen die Objekte oder Vorgänge auf Bild- oder Videomaterial nicht gesamtheitlich bzw. in einem Sinnzusammenhang erfassen,

sondern lediglich aufgrund der algorithmischen Identifikation und Korrelation von verschiedenen Pixelpunkten. Dies bedeutet, dass kontextuales Hintergrundwissen oder Assoziationen zumindest heute noch nicht in die Bilderkennung miteinfließen.

Darüber hinaus kann das bewusste oder zufällige Einspeisen von kleinsten Veränderungen oder Manipulationen wie beispielsweise einzelnen Farbpunkten dazu führen, dass ein Bilderkennungsprogramm zu fehlerhaften Resultaten kommt. Das Massachusetts Institute of Technology publiziert verschiedene Fälle von sogenannten Adversarial Effects (zu Deutsch: widersprüchliche Wirkungen).³ Durch das Hinzufügen von einzelnen Pixeln, die für das menschliche Auge unsichtbar sind, konnte ein Bilderkennungsprogramm Pandas nicht mehr zuverlässig als solche erkennen. Der zweite Fall demonstriert, wie diese Effekte auch bei praktischen Anwendungen die Ergebnisse entscheidend beeinflussen können: Das Anbringen eines Aufklebers auf einem Stoppsignal führte dazu, dass ein autonomes Fahrzeug ein Verkehrsschild fälschlicherweise als Geschwindigkeitsbegrenzung identifizierte. Wichtig ist es für die Forschung zu klären, ob diese ungewollten Effekte eine untergeordnete Problematik darstellen oder ob sie eine fundamentale Schwäche der aktuell verwendeten Bilderkennungsmethoden offenlegen.⁴

Embodiment: Wahrnehmung durch Verkörperung

Ein zentrales Ziel von KI-Systemen ist die Wahrnehmung und Interpretation ihrer Umgebung als Entscheidungsgrundlage. Das sogenannte Embodiment (zu Deutsch: Verkörperung) geht davon aus, dass richtige Wahrnehmung nicht allein auf einer Abbildung von sensorischen Signalen beruhen kann, sondern eine physische Interaktion mit der Umwelt voraussetzt.⁵ Im Kontext der Künstlichen Intelligenz bedeutet die These, dass die Konstruktion von wahrnehmenden und intelligenten Maschinen eine körperliche Wechselwirkung mit ihrer Umgebung erfordert.⁶ Bestehende Konzepte der KI-Forschung, die ihr Hauptaugenmerk auf die Rekonstruktion menschlicher Gehirnfunktionen richten, müssten folglich um eine körperliche Dimension ergänzt werden, um dieser Konzeption von Kognition zu entsprechen.

Natural Language Processing: Simulation von Sprachfähigkeit

KI-Systeme können einen gewissen Grad an Sprachfähigkeit simulieren, indem Algorithmen innerhalb von Ton- oder Textdateien Regeln identifizieren und diese anwenden, um eine Interaktion zwischen Menschen und Maschinen zu ermöglichen. Im interdisziplinären Forschungsfeld Natural Language Processing (zu Deutsch: Computerlinguistik) arbeitet die Sprachwissenschaft mit der Informatik gemeinsam an der Analyse, Interpretation und Generierung von natürlicher Sprache. Da gesprochene bzw. geschriebene Konversationen und Aussagen als Grundlage dienen, verwenden die Computersysteme in den meisten Fällen unstrukturierte Daten. Um Sprache von der Schallinformation bis hin zur Bedeutung zu analysieren, werden verschiedene Techniken und Methoden miteinander kombiniert.⁷ Die phonetische Ebene ist lediglich für die gesprochene Sprache relevant, nicht aber für reine Textverarbeitung. In einem ersten Schritt werden innerhalb von Audiosignalen durch akustische Mustererkennung Lauteinheiten aufgespürt und klassifiziert. Die eingegangenen Signale werden mit einem Wörterbuch und einem Sprachmodell abgeglichen, um die Wahrscheinlichkeiten von bestimmten Laut- bzw. Wortkombinationen zu berechnen. Nachdem die Audiosignale in Textform umgewandelt wurden, können die Buchstabenketten mittels Tokenisierung in Worteinheiten aufgeteilt werden. Dabei gibt es zur syntaktischen Verarbeitung diverse zusätzliche Schritte, wie beispielsweise die morphologische

Segmentierung, bei der zur Ableitung von grammatischen Informationen Fallmarkierungen, Pluralformen oder Wortzusammensetzungen interpretiert werden. Durch Semantik- und Dialoganalysen können schliesslich die Bedeutungen und Beziehungen zwischen Wörtern und Sätzen abgebildet werden. Spezifische Verfahren zur Erkennung von Eigennamen können dabei helfen, die Klassifikation von Personen, Firmen oder Orten zu automatisieren. Im Forschungsfeld Natural Language Processing werden – vergleichbar zum maschinellen Sehen – einzelne Sprachkomponenten analysiert und in das Gesamtsystem der Sprache übertragen. Auf diesen Prinzipien beruht die Spracherkennung bei kommunikativen Assistenzsystemen wie Siri als Steuerungssoftware von Apple-Produkten oder Alexa von Amazon zur Interaktion mit digitalen Anwendungen oder der Smart-Home-Infrastruktur.

Neben der Analyse von Schallinformationen können digitale Systeme auch bestehende Textdateien verarbeiten. Text Mining ermöglicht es, innerhalb von unstrukturierten Textdaten automatisiert Bedeutungsstrukturen herauszuarbeiten. Das Verfahren eignet sich insbesondere zur Extraktion von relevanten Informationen und Aussagen aus grossen Textdateien. Text Mining erlaubt es, mithilfe von Algorithmen inhaltliche Hypothesen zu generieren und diese automatisiert zu überprüfen.⁸ Ein weitverbreitetes Verfahren zur Auswertung von Textinhalten ist die Sentimentanalyse, bei der versucht wird, beispielsweise einzelne Kommentare in den sozialen Medien auf ei-

«Anwendungen zur Sprach- erkennung und -übersetzung sind soweit ausgereift, dass sie in unseren Alltag Einzug halten.»

Jana Koehler, Informatikprofessorin und wiss.-techn.
Geschäftsführerin des Deutschen Forschungszentrums für KI

ner Polaritätsachse zwischen positiv und negativ einzuordnen. Dadurch können subjektive Meinungen ausgewertet werden, um Rückschlüsse über die Popularität von Zeitungsbeiträgen, Produkten oder Personen zu ziehen.

Die Potenziale der Sprach- und Textverarbeitung werden allerdings durch mehrere limitierende Faktoren beschränkt, wie beispielsweise durch die Unzuverlässigkeit bei der Übersetzung in unterschiedliche Sprachen oder der mangelhaften Erkennung von seltenen Dialekten sowie der Interpretation von rhetorischen Stilmitteln wie Ironie oder Metaphern. In der Konsequenz fehlt es der Sprachverarbeitung oftmals an einer Einbettung von Wörtern, Fragen, Aussagen oder Dialogen in einen erweiterten Sinnzusammenhang.

Die automatisierte Mustererkennung kommt jedoch nicht nur bei der Nachahmung von menschlichen Sinneswahrnehmungen zum Einsatz. Prinzipiell können alle digitalisierbaren Informationen oder beste-

hende Datensätze durch Mustererkennung analysiert und klassifiziert werden. Dies gilt für meteorologische Daten ebenso wie für Aktienkurse oder genetische Informationen von Menschen. Das Gebiet der Mustererkennung ist damit äusserst breit und reicht von der Analyse von Stimmen, Handschriften, DNA, Speichel und Atem, Gesichtern und Produkten über die Auswertung von Textdokumenten und Videodateien bis hin zur Erkennung von Objekten im Strassenverkehr oder im Weltraum.



WIE MASCHINEN LERNEN

Lernfähigkeit ist wie das Klassifizieren ein integraler Bestandteil von intelligentem Verhalten. So lässt sich wissenschaftlich eine klare Korrelation zwischen klassischen Intelligenztests und der Lernfähigkeit von Menschen nachweisen.⁹ Dieser Zusammenhang beruht insbesondere darauf, dass durch Lernen der Aufbau einer Wissensbasis und das Aneignen von flexiblen Problemlösungsansätzen gefördert werden. Lernen ist dabei ein äusserst vielfältiges Phänomen und wird nicht nur bei relativ simplen Alltagsvorgängen wie dem Merken eines neuen Namens verwendet, sondern auch bei hochkomplexen Aktivitäten wie der Herleitung von physikalischen Zusammenhängen wie im Fall von Albert Einsteins Relativitätstheorie.

Im Kontext der Künstlichen Intelligenz beschreibt Lernen die Fähigkeit, durch vergangene Beobachtungen den Erfolg beim Lösen von künftigen Aufgaben zu verbessern. Dabei gibt es Aufgaben, die von Menschen auf sehr intuitive Art und Weise und

ohne Anstrengung gelöst werden können, bei denen die meisten Computersysteme jedoch versagen. Ein Beispiel hierfür ist das Erkennen der Gesichter von Familienangehörigen oder Freunden. Regelbasierte Computersysteme scheitern an der Gesichtserkennung, weil die Variabilität von Gesichtsausdrücken, Kopfstellungen, Lichteinflüssen oder Altersunterschieden nur äusserst schwer durch vorgegebene Zusammenhänge zu erfassen ist.

Gerade bei für Menschen intuitiv lösbaren Problemen ist die Verwendung von Machine Learning (zu Deutsch: maschinelles Lernen) nach dem heutigen Forschungsstand eine der wenigen Methoden, die verlässliche Resultate erzielt. Machine Learning ist für weite Teile des Gebiets der Künstlichen Intelligenz von zentraler Bedeutung. Entscheidend dabei: Computersysteme erlangen zur Lösung von Aufgaben eine Lernfähigkeit, ohne explizit dafür programmiert zu werden.¹⁰ Sie optimieren sich vielmehr autonom und basierend auf Trainingsdaten, erkennen Gesetzmässigkeiten und Muster und verallgemeinern diese durch Inferenz. So lassen sich auch komplexe Aufgaben mit hoher Variabilität, wie beispielsweise die zuvor beschriebene Bilderkennung oder Sprachverarbeitung, zuverlässiger lösen, weil nicht sämtliche Regeln vorgegeben werden müssen. Wie beim Menschen gibt es auch bei den Maschinen verschiedene Arten des Lernens. Grundsätzlich kann zwischen den drei Methoden Supervised, Unsupervised und Reinforcement Learning unterschieden werden.

Supervised Machine Learning:

Training durch vorgegebene Beispiele

Beim Supervised Learning (zu Deutsch: überwachtes Lernen) wird ein Programm mithilfe von vorgegebenen Beispielen trainiert, um bei neuen Aufgaben zu möglichst genauen Schlussfolgerungen zu gelangen. Dabei werden Datenlabels verwendet, die dem System eindeutig vorgeben, nach welchen Mustern und Kategorien gesucht werden soll. Dementsprechend nimmt der Mensch als Überwacher und Wegweiser beim Training der Modelle eine zentrale Rolle ein, indem er die Datensätze aufbereitet und den Trainingsprozess kontrolliert. In der medizinischen Diagnostik kann Supervised Learning beispielsweise verwendet werden, um auf Bildern automatisiert Tumore zu erkennen und diese vorgegebenen Kategorien zuzuordnen. Dabei stützen sich die Trainingsbeispiele auf das Fachwissen und die Einschätzungen von Expertinnen und Experten, die verschiedene Tumorarten basierend auf ihrer medizinischen Expertise unterscheiden und dadurch den Lernprozess und die Zuverlässigkeit von bildgebenden Verfahren unterstützen.¹¹

Unsupervised Machine Learning:

lernen ohne Instruktionen

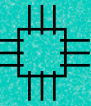
Beim Unsupervised Machine Learning (zu Deutsch: unüberwachtes Lernen) sucht das Programm eigenständig nach Mustern innerhalb eines Datensatzes, ohne zuvor konkrete Instruktionen zu erhalten. Der Algorithmus ist dann in der Lage, ohne Trainingsbeispiele oder Datenlabels verschiedene Arten vorzuschlagen, wie der Datensatz aufgrund seiner internen Zu-

sammenhänge kategorisiert werden könnte. Ein Verfahren, das dazu verwendet wird, ist die Clusteranalyse, bei der innerhalb des Datensatzes Ähnlichkeitsstrukturen identifiziert werden, um Gruppen von Analyseobjekten mit ähnlichen Eigenschaften zu bilden. Diese Methode kommt oftmals zum Einsatz, wenn noch nicht feststeht, nach welchen Zusammenhängen gesucht werden soll. Neben der Kategorisierung und Visualisierung von Datensätzen ist Unsupervised Machine Learning vor allem für die Detektion von Unregelmässigkeiten geeignet. Konkrete Anwendungsfälle sind die Aufdeckung von Anomalien bei Kreditkartentransaktionen oder die frühzeitige Erkennung von Cyberangriffen, wobei beispielsweise ungewöhnliche Zugriffsversuche auf Datenspeicher signalisiert werden.¹²

Reinforcement Machine Learning:

lernen durch Iterationen

Beim Reinforcement Learning (zu Deutsch: bestärkendes Lernen) versucht das System, ein klar definiertes Ziel mit einer Trial-and-Error-Strategie zu erreichen und testet dazu verschiedene Varianten. Diejenigen Varianten, die funktionieren, werden belohnt, wohingegen schlechte Entscheidungen bestraft werden. Dies wird mithilfe von mathematischen Belohnungs- bzw. Bestrafungsfunktionen erreicht, die dem System vorgeben, welcher Endzustand erreicht werden soll. Die Methode orientiert sich stärker am Lernverhalten des Menschen, da sie auf praktischer Erfahrung und wiederholten Iterationen mit der Umwelt beruht. Reinforcement Learning ist



«Die Kombination von riesigen Datenmengen und erhöhten Rechenkapazitäten erlaubt es Künstlichen Neuronalen Netzwerken, sehr komplexe Zusammenhänge zu erlernen.»

Thomas Hofmann, Professor für Machine Learning an der ETH Zürich

die neuste Generation von Machine Learning und hat in den letzten Jahren nachhaltige Erfolge ermöglicht. Im Jahr 2017 hat Googles KI-Sparte Deep Mind ein System entwickelt, das innerhalb von nur 70 Stunden ein hohes Leistungsniveau im strategischen Brettspiel Go erreicht hat.¹³ Beim Start wurden dem System keinerlei Kenntnisse des Spiels zur Verfügung gestellt – ausser den Regeln. Das Programm hat sich eigenständig basierend auf Reinforcement Learning Schritt für Schritt verbessert, indem es mehrere Millionen Partien gegen sich selbst ausgetragen hat. Dadurch konnte es den weltbesten Spieler mehrmals hintereinander besiegen. In Zukunft dürfte Reinforcement Learning nicht nur bei strategischen Spielen wie Go oder Schach, sondern auch vermehrt bei Alltagsanwendungen zum Einsatz kommen. Eine der Schwierigkeiten besteht jedoch darin, dass in der realen Welt beispielsweise bei der Auswahl von Therapievorschlügen für Patienten oder der Steuerung von autonomen Fahrzeugen anders als bei virtuellen Simulationen nicht immer Tausende von Iterationen durchgeführt werden kön-

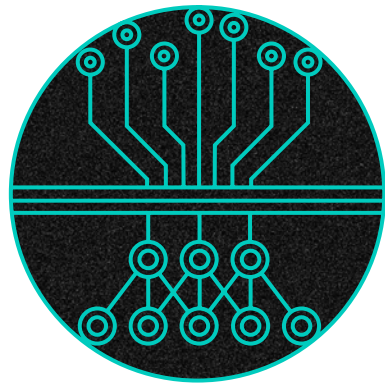
nen und sich die Rahmenbedingungen kontinuierlich ändern. Des Weiteren verlieren auch die Systementwickler in manchen Fällen die Kontrolle über den Lernprozess. Als Facebook im Jahr 2017 durch Reinforcement Learning einen sprachbasierten Software-Roboter mit Verhandlungskompetenzen entwickeln wollte, hat dieser nach Tausenden Iterationen eine eigene, für Menschen unverständliche Form von Englisch entwickelt, die auch für die Programmierer nicht mehr nachvollziehbar war.¹⁴

Aufschwung durch Künstliche Neuronale Netzwerke

Einen Aufschwung erlebte Machine Learning nach der Jahrtausendwende auch deshalb, weil Künstliche Neuronale Netzwerke erfolgreich weiterentwickelt werden konnten. Die grundlegende Idee, Gehirnaktivitäten sowohl in mathematischen Modellen als auch in der Informatik nachzubilden, ist dabei so alt wie das Gebiet der Künstlichen Intelligenz selbst. In diesem Kontext geht es weniger darum, biologische Gehirnstrukturen nachzubauen, sondern auf einer abstrakten Ebene die Me-

chanismen von neuronalen Netzwerken zu imitieren. Lange Zeit fehlte der Forschung die nötige Rechenleistung, um komplexe Berechnungen basierend auf künstlichen Neuronen durchzuführen. Erst durch die Verwendung von Graphics Processing Units (GPU), die ursprünglich als Grafikkarten für Videospiele entwickelt wurden, konnten maschinelle Lernverfahren basierend auf Künstlichen Neuronalen Netzwerken immer komplexere Probleme lösen. Die künstlichen Neuronen können dabei Signale verarbeiten und an verbundene Neuronen weiterleiten. Die Netzwerke verbessern beispielsweise ihre Ergebnisse, indem sie neue Verbindungen zwischen künstlichen Neuronen aufbauen, bestehende Verbindungen löschen oder anders gewichten, oder indem sie die Schwellenwerte für die Signalübertragung anpassen. Die Methode wird als Deep Learning bezeichnet, sobald mehrere Zwischenlagen von Künstlichen Neuronalen Netzwerken verwendet werden. Der Begriff Deep kommt daher, dass unterschiedliche Schichten von Rechenknoten zusammenwirken, um Daten zu durchsuchen und ein optimiertes Endergebnis zu liefern. Deep Learning kann mit allen drei zuvor beschriebenen Arten des Lernens kombiniert werden und verspricht insbesondere bei grossen Datenmengen eine erhöhte Genauigkeit gegenüber herkömmlichen Methoden. Künstliche Neuronale Netzwerke werden beispielsweise in der Medizin zur genaueren Diagnose von Lungenknoten bei bildgebenden Verfahren oder zur schnelleren und zuverlässigeren maschinellen Übersetzung von seltenen Wortkombinationen verwendet.^{15/16}

Machine-Learning-Methoden werden immer mehr zum Kern von Anwendungen mit Künstlicher Intelligenz. Unter Forschern und Forscherinnen ist der Ausdruck Künstliche Intelligenz denn auch umstritten. Da die meisten der heutigen Systeme in der einen oder anderen Form auf Machine-Learning-Methoden beruhen, bevorzugen viele Forscher und Forscherinnen diesen Ausdruck gegenüber dem verbreiteten Sammelbegriff der Künstlichen Intelligenz. Neben Machine Learning gibt es allerdings weitere relevante Verfahren zur Optimierung von Datenanalysen. Die Entwicklung von evolutionären Algorithmen orientiert sich beispielsweise an evolutionsbiologischen Prinzipien.¹⁷ Zufälle und Mutationen werden dabei zur Lösung von spezifischen Problemstellungen verwendet. Am Anfang steht die willkürliche Generierung von unzähligen Varianten verschiedener Codezeilen. Danach werden die unterschiedlichen Versionen getestet, um zu überprüfen, ob sie das vorgegebene Ziel erreichen. Falls eine Version die Aufgabe zufälligerweise marginal besser löst als die anderen, werden leicht adaptierte Kopien erstellt, die dann wiederum gegeneinander antreten. Dieser Prozess wird so lange fortgeführt, bis das ursprüngliche Problem zuverlässig gelöst werden kann. Ansätze wie die Entwicklung von evolutionären Algorithmen zeigen, dass neben Machine Learning auch andere Optimierungsverfahren die Zukunft der Künstlichen Intelligenz bestimmen könnten.



WO INTELLIGENTE MASCHINEN AN GRENZEN STOSSEN

Um das mittel- und langfristige Potenzial Künstlicher Intelligenz realistisch einschätzen zu können, gilt es auch die Limitationen zu beurteilen. Viele der Ankündigungen zur Leistungsfähigkeit und zu den Potenzialen künstlich intelligenter Systeme der letzten Monate und Jahre dürften darauf beruhen, dass kritische Aspekte nicht oder zu wenig berücksichtigt wurden. Dies zeigt sich auch daran, dass einige Anwendungen fälschlicherweise als Künstliche Intelligenz beschrieben wurden. Oftmals handelte es sich vielmehr um einfache Algorithmen, die mit hoher Rechenleistung gute Ergebnisse erzielten, jedoch nicht über autonome Lernfähigkeiten verfügten. Mit Blick auf die relevanten Einschränkungen gilt es verschiedene Aspekte zu berücksichtigen, die über die mangelnde Lernfähigkeit hinausgehen.

Kausale Inferenz: Korrelation ist nicht gleich Kausalität

Machine-Learning-Modelle identifizieren innerhalb von Trainingsdaten automatisiert Korrelationen. Die Anzahl der Korrelationen, die von den selbstlernenden Algorithmen berücksichtigt werden, steigt mit zunehmender Grösse des Datensatzes stark an. Doch längst nicht jeder statistisch signifikante Zusammenhang steht auch mit einer Kausalität in Verbindung. Die heutigen KI-Systeme haben keine Möglichkeit, die identifizierten Regelmässigkeiten auf ihren Gehalt und ihre Bedeutung für die reale Welt jenseits von statistischen Modellen zu überprüfen. Deshalb besteht die Gefahr, durch den Einsatz von Machine-Learning-Methoden falsche Rückschlüsse zu ziehen. Nämlich solche, die nicht auf einer kausalen Verbindung aufbauen. Die sogenannte kausale Inferenz ist deshalb im Kontext von Machine Learning ein aktueller Forschungsschwerpunkt, um durch den Einsatz von selbstlernenden Algorithmen die Komplexität der kausalen Mechanismen der realen Welt abzubilden.¹⁸

Overfitting: mehr Informationen führen nicht zu besseren Ergebnissen

Eine Herausforderung für Künstliche Intelligenz beruht darauf, dass sich die Modelle zur Mustererkennung stark an die verwendeten Trainingsdaten anpassen. Beim sogenannten Overfitting (zu Deutsch: Überanpassung) werden bei der Anwendung eines Modells auf einen neuen Datensatz unter Umständen falsche Rückschlüsse gezogen. Der Grippedienst von Google, der bereits kurz nach seiner Einführung auf-

«Aktuelle KI-Ansätze beruhen auf statistischen Verfahren und automatisierter Mustererkennung. Mit der menschlichen Intelligenz hat dies nur sehr wenig zu tun.»

Pascal Kaufmann, Neurowissenschaftler und KI-Unternehmer

grund von fehlerhaften Angaben eingestellt werden musste, veranschaulicht die Problematik exemplarisch.¹⁹ Die folgende Idee lag dem Dienst zugrunde: Sobald Menschen erste Grippe-symptome aufweisen, recherchieren sie via Google-Suche nach Informationen und signalisieren nahezu in Echtzeit das Auftreten von Grippefällen in einer bestimmten Region. Das Versprechen lag darin, Grippeepidemien noch vor dem US-Gesundheitsdienst aufgrund der Suchanfragen aufzudecken. Es stellte sich jedoch heraus, dass innerhalb der immensen Trainingsdaten auch völlig zufällig korrelierte Suchanfragen wie «High School Basketball» zur Mustererkennung berücksichtigt wurden. Die Fehleranfälligkeit des Modells im Rahmen einer fatalen Grippe-welle im Jahr 2013 resultierte in der Einstellung des Programms.

Mangelnde Übertragbarkeit: Modelle sind kontextspezifisch

Machine-Learning-Modelle lassen sich oftmals nur für sehr spezifische Aufgaben verwenden und sind deshalb nur beschränkt auf verwandte Problemstellungen übertragbar. Dies führt vor allem in der Praxis zu

grossen Schwierigkeiten, sobald versucht wird, ähnliche, aber nicht identische Aufgaben zu lösen. Konkret kann dies bedeuten, dass ein Modell zur autonomen Steuerung von Fahrzeugen, das anhand von amerikanischen Stassenverkehrsdaten trainiert wurde, in einem europäischen Kontext nicht zuverlässig funktioniert. Auch Modelle zur Spracherkennung lassen sich beispielsweise nicht immer von einer Sprachfamilie auf eine andere übertragen. So haben Spracherkennungsprogramme deutlich mehr Probleme mit tonalen Sprachen wie Kantonesisch, bei denen die Tonhöhe für die Unterscheidung von Wortbedeutungen verwendet wird.²⁰ Die unterschiedlichen Anforderungen spielen auch in der Bilderkennung eine Rolle. Die US-Regierung hat bei einem Vergleich von Algorithmen zur Gesichtserkennung festgestellt, dass auch die leistungsfähigsten Programme fehleranfälliger werden, wenn die Gesichter eine dunkle Hautfarbe haben.²¹ Über die Gründe für die geringere Zuverlässigkeit im Vergleich zur Erkennung von Gesichtern mit helleren Hauttypen ist sich die Forschung nach wie vor uneins. Die unterschiedlichen Bedingungen und Anfor-

derungen je nach Anwendungsbereich erhöhen die Komplexität der Mustererkennung jedenfalls stark.

Dark Data: fehlende Daten und mangelnde Datenqualität

Eine der grössten Schwierigkeiten in der Entwicklung von Machine-Learning-Methoden liegt in der Beschaffung von hochwertigen und aktuellen Daten. Um wissenschaftlichen Ansprüchen zu genügen und optimale Ergebnisse zu erreichen, setzen manche Methoden eine aufwendige Bereinigung und Kennzeichnung der Datensätze voraus. Im Gesundheitsbereich haben Unternehmen wie IBM das Problem der Verfügbarkeit von hochwertigen Daten beispielsweise unterschätzt. Seit 2011 treibt der amerikanische Konzern seine KI-Sparte Watson im Gesundheitssektor voran und ist seither über 50 Partnerschaften eingegangen. Bei Kunden und Kundinnen, Investoren und Investorinnen wurden sehr hohe Erwartungen an die Leistungsfähigkeit von KI-Anwendungen wie der algorithmischen Genomanalyse oder der personalisierten Diabetestherapie geweckt. Doch nur die wenigsten Partnerschaften waren kommerziell erfolgreich. Einer der Hauptgründe für die ernüchternde Bilanz lag im unzureichenden Zugang zu qualitativ hochwertigen Daten.²² Ein Beispiel aus der Bilderkennung illustriert die hohen Anforderungen. So benötigt ein Deep-Learning-Programm heute 60000 gekennzeichnete Trainingsbeispiele, um Bilder von Fahrzeugen und Tieren aus zehn verschiedenen Kategorien zu unterscheiden und einander zuzuordnen.²³ Bei komplexeren Problemen

kann die erforderliche Stichprobengrösse auch bis zu einer Million ansteigen. Deshalb zielen aktuelle Forschungsambitionen darauf ab, auch aus kleineren Datensätzen zu lernen. Beim sogenannten Transfer Learning wird versucht, das erlangte Wissen aus der Bearbeitung eines Problemfelds auf neue, aber verwandte Problemfelder mit schlechterer Datenverfügbarkeit zu übertragen. So können Lernmodelle beispielsweise zur Unterscheidung von Hunden und Katzen dazu verwendet werden, um andere Tierarten wie Wespen oder Ameisen zu kategorisieren.²⁴

Zentralisierung der Daten: Konzentration bei Plattformanbietern

Herkömmliche Machine-Learning-Methoden setzen auf eine Zentralisierung der Trainingsdaten in einem leistungsfähigen Rechner oder einem Datencenter. Dies ist in der Öffentlichkeit mit erheblichen Datenschutzbedenken verbunden. Deshalb wird aktuell intensiv an einem neuen Ansatz, dem sogenannten Federated Learning, geforscht.²⁵ Die grundlegende Idee besteht darin, dass sich verschiedene Endgeräte wie beispielsweise Mobiltelefone ein gemeinsames Machine-Learning-Modell teilen, ohne dass hierfür persönliche Daten von einem Nutzer oder einer Nutzerin an einen zentralen Anbieter abwandern. Das individuelle Nutzergerät lädt dafür das aktuelle Lernmodell herunter und verbessert aufgrund der persönlichen Daten die Vorhersagen. Die Verbesserungsvorschläge werden als Änderungen des Modells zusammengefasst und in verschlüsselter Form an den Anbieter zurückgesendet. Der

entscheidende Unterschied: Sämtliche Nutzerdaten bleiben auf dem jeweiligen Endgerät. Dieser Ansatz könnte in Zukunft dazu führen, dass die Fähigkeit, Machine-Learning-Methoden zu verwenden, von der Notwendigkeit einer zentralen Datenspeicherung entkoppelt wird. Gerade in einem Sektor wie dem Gesundheitsbereich, in dem es um sensitive und personenbezogene Daten geht, könnten somit Datenschutzbedenken zumindest teilweise adressiert werden.

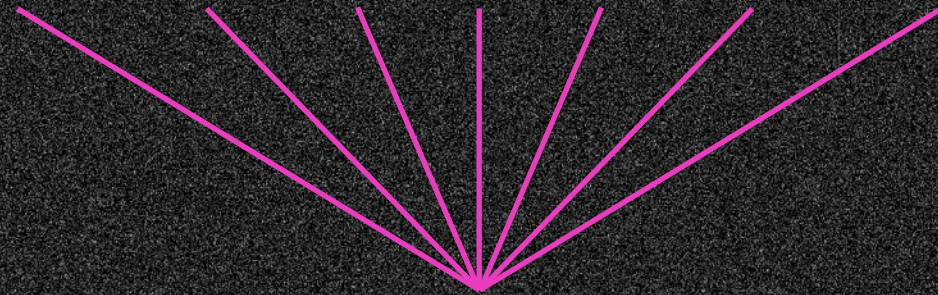
Black-Box-Effekte: Lösungen ohne erklärbare Herleitung

Ein wichtiges Ziel der KI-Forschung ist es, kurz- bis mittelfristig das Problem der Erklärbarkeit von Machine-Learning-Systemen zu entschärfen. Gerade bei Aufgaben, bei denen eine Vielzahl von Einflussfaktoren eine Rolle spielen, ist es oftmals unklar, wie ein System zu einem bestimmten Resultat gekommen ist. Ein KI-System mit unnachvollziehbaren Ergebnissen wird deshalb auch als Black Box bezeichnet. Das Feld der erklärbaren Künstlichen Intelligenz beschäftigt sich mit der Frage, wie die Verfahren von dynamischen und nicht linear programmierten Systemen wie Künstlichen Neuronalen Netzwerken auf eine für Menschen nachvollziehbare Art dargestellt werden können. Einige Forscher und Forscherinnen argumentieren, dass eine geringere Zuverlässigkeit der Ergebnisse in Kauf genommen werden muss, um die Nachvollziehbarkeit der Resultate zu verbessern.²⁶ Um diese Problematik zu entschärfen, werden gezielt Verfahren wie die kontrafaktische Methode entwickelt, bei der versucht

wird, aufgrund von veränderten Inputdaten die Auswirkungen auf die Endergebnisse von Künstlichen Neuronalen Netzwerken zu analysieren, um dadurch Rückschlüsse über die Funktionsweisen der Algorithmen zu ziehen.²⁷

Benötigte Rechenleistung: hohe Kosten und Umweltauswirkungen

Aus ressourcentechnischer Sicht stellt sich neben dem Zugang und der Verarbeitung von Daten eine weitere Herausforderung. Mit der hohen Rechenleistung, die zum Training von Lernmodellen benötigt wird, steigt auch der Energiebedarf. So haben Forscher der University of Massachusetts berechnet, dass der vollständige Trainingszyklus eines weitverbreiteten Modells, das zur Sprachverarbeitung verwendet wird, bis zu fünfmal so hohe CO₂-Emissionen wie ein durchschnittliches Auto über seine ganze Lebensdauer verursacht.²⁸ Die effektiven Emissionen hängen dabei stark vom jeweiligen Energiemix ab, der zum Training der Modelle verwendet wird. Dies ist einer der Gründe, weshalb grosse Technologiekonzerne wie Google und Amazon vermehrt hohe Investitionen in erneuerbare Energien tätigen. Darüber hinaus ist die immense Rechenleistung, die zum Training der Lernmodelle benötigt wird, auch aus wirtschaftlicher Perspektive mit sehr hohen Kosten verbunden.

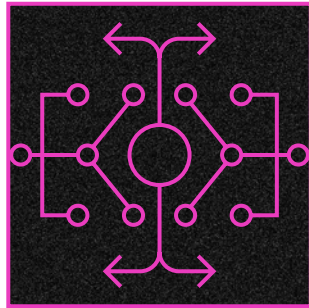


3

Anwendungsbereiche

WAS MASCHINEN KÖNNEN

Mustererkennung und Lernfähigkeit bilden das Fundament der Anwendung von KI-Technologien. Doch eine Innovation entsteht erst, wenn Anwendungen auch einen Nutzen bringen. Fünf Anwendungsbereiche stehen für den heutigen und künftigen Praxiseinsatz von Künstlicher Intelligenz im Zentrum: entscheiden, vorhersagen, kommunizieren, Bewegungen steuern und generieren. Im Unterschied zu traditionellen Algorithmen wird der Mehrwert der lernenden Systeme darin bestehen, dass sie eigenständig auf bisher unbekannte Herausforderungen reagieren können und dabei einen grösseren Teil des Aufgabenspektrums in Organisationen, aber auch im Alltag von Menschen übernehmen. Die nahe und auch fernere Zukunft werden entsprechend durch mehr Interaktionen mit künstlich intelligenten Systemen geprägt sein, die als Entscheidungshelfer mit besseren Prognosen unterstützen, die aber auch eigenständig einfache Inhalte erarbeiten werden. Gerade durch die automatisierte Generierung von neuen Lösungsansätzen von der Medizin bis zur Architektur oder insbesondere durch die virtuelle Simulation von Menschen wird eine neue Dimension der digitalen Transformation angestoßen, die über das hinausgeht, was bislang möglich war.



ENTSCHEIDEN

Durch automatisierte Mustererkennung und Lernfähigkeit können KI-Systeme Datensätze interpretieren und dadurch Empfehlungen bzw. Entscheidungen ableiten. Insbesondere in Situationen mit einer Vielzahl von Optionen können KI-Systeme mit selbstlernenden Algorithmen optimierte Ergebnisse erzielen. Im Folgenden werden einzelne Varianten vorgestellt, wie KI-Systeme Empfehlungen abgeben oder Entscheidungen fällen.

Assistenz durch Empfehlungen

KI-basierte Empfehlungssysteme können bei medizinischen Diagnosen, Kaufentscheidungen oder bei der Routenplanung im Verkehr unterstützen. Wissenschaftler der UC San Francisco haben basierend auf Deep-Learning-Methoden ein KI-System entwickelt, das Ultraschallkardiografien – bestehend aus Video- und Standbildern – in 92% der Fälle richtig klassifizieren konnte.¹ Erfahrene Kardiologen kamen hingegen lediglich auf einen Genauigkeitswert von 79%. Die Unterstützung von medizinischen Diagnosen bei bildgebenden Verfahren ist

auch in der Radiologie und Dermatologie von Bedeutung, wobei die finale Entscheidung heute nach wie vor meist von einem Menschen getroffen wird. Darüber hinaus stossen KI-basierte Diagnosesysteme bei höherem Komplexitätsgrad der Symptome an Grenzen. So sinkt die Genauigkeit der Resultate bei Patienten mit verschiedenen, gleichzeitig auftretenden Krankheitsbildern deutlich.

Im Strassenverkehr optimieren KI-Systeme bereits heute menschliche Entscheidungen, indem unter einer grossen Anzahl von Navigationsmöglichkeiten die schnellste Route berechnet wird. Basierend auf Echtzeitdaten werden die empfohlenen Navigationsentscheidungen fortlaufend angepasst. Auch wenn die Entscheidung über den eingeschlagenen Weg aktuell noch beim menschlichen Fahrer liegt, ersetzen digitale Assistenzsysteme zunehmend menschliche Erfahrungswerte und Intuition.

Autonome Entscheidungen

In bestimmten Fällen treffen KI-Systeme bereits heute autonom Entscheidungen und führen ohne menschliche Intervention Aktionen aus. Selbstlernende Algorithmen kommen etwa an den Finanzmärkten zum Einsatz, wo sie für institutionelle Anleger Börsenwerte automatisiert kaufen bzw. verkaufen. Bei der Rekrutierung von Arbeitskräften können Bewerberinnen und Bewerber durch KI-Systeme aufgrund von automatisiert ausgewerteten Videointerviews vorselektioniert werden.² In den sozialen Medien entscheiden KI-Algorithmen darüber, welche Inhalte einem be-

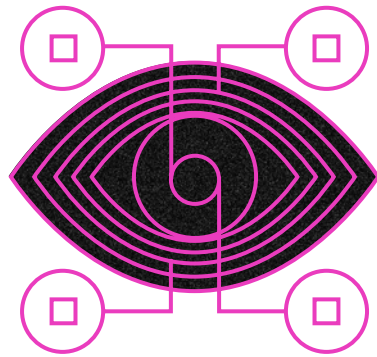
«Informatiker, die KI-basierte Entscheidungssysteme programmieren, sollten in der Entwicklung möglichst früh auch ethische Fragestellungen berücksichtigen.»

Eva Weber-Guskar, Philosophieprofessorin und
Fellow der Digital Society Initiative der Universität Zürich

stimmten Nutzer gezeigt werden, um dadurch die Verweildauer auf den Plattformen zu maximieren. Es bleibt zu berücksichtigen, dass die Empfehlungen und Entscheidungen zwar neu in dem Sinn sind, dass sie jeweils situativ berechnet werden. Allerdings sind die verwendeten Parameter ein limitierender Faktor für die Bandbreite von Ergebnissen: Entsprechend sind etwa Musik-Streaming-Plattformen nur in der Lage, jene Musik zu empfehlen, die sich aus dem Musikgeschmack der Vergangenheit ableiten lässt.

ANWENDUNGEN HEUTE UND MORGEN

- **Retail:** Bei der dynamischen Preisgestaltung (Dynamic Pricing) von Lebensmitteln am Verkaufspunkt können unter anderem basierend auf Wetterdaten, Lagerbestand und Haltbarkeitsdaten die Verkaufspreise fortlaufend angepasst werden.
- **Medien:** Identifikation von Nutzerinteressen anhand der Sehgewohnheiten auf Video-on-Demand-Plattformen dienen als Grundlage für die Produktion neuer Filme und Serien.
- **Energie:** Intelligente Energienetze, sogenannte Smart Grids, optimieren dezentrale Energieproduktion, -speicherung und -verbrauch.
- **Finanzsektor:** Roboadvisor personalisieren und optimieren Anlageentscheidungen basierend auf individuellen Vorgaben zu Risikoeinstellungen oder Themenschwerpunkten.
- **HR:** Automatisierte Identifikation von Mitarbeiterkompetenzen unterstützen den Aufbau von Wissensnetzwerken und das Recruiting.



VORHERSAGEN

Nebst den Fähigkeiten, Entscheidungen zu fällen und Empfehlungen abzugeben, können KI-Systeme Voraussagen treffen. Aufbauend auf der Analyse von historischen Daten kann der Algorithmus für neue Datensätze eine statistische Wahrscheinlichkeit für das Eintreten künftiger Zustände berechnen. Machine-Learning-Methoden verbessern die Modelle in mehrstufigen Verfahren, um die Zusammenhänge, Trends und Korrelationen immer genauer herauszuarbeiten.

Prognosen durch Klassifizierung und Regression

Die Art der Anwendung bestimmt dabei die Prognosemethode. Soll etwa die Temperatur eines Motors zu einem künftigen Zeitpunkt berechnet werden, kommen sogenannte Regressionsmodelle zum Einsatz, die basierend auf historischen Daten eine Funktion ableiten, mit der die Temperaturentwicklung des Systems bestimmt werden kann. Wenn der KI-Algorithmus lediglich bestimmen soll, ob der Motor wahrscheinlich überhitzt oder nicht, werden die

«Überall dort, wo wir Vorhersagen aufgrund von Erfahrungswerten treffen, kann Künstliche Intelligenz auf eine riesige Wissensbasis zurückgreifen.»

Thomas Hofmann, Professor für Machine Learning an der ETH Zürich

historischen Daten mit der sogenannten Klassifizierung in zwei Gruppen eingeteilt und neue Werte entsprechend zugeordnet. Seit 2013 kommt etwa in Floridas Justizsystem ein Algorithmus zum Einsatz, der basierend auf 137 Datenpunkten die Wahrscheinlichkeit bestimmt, ob ein Gefangener wieder straffällig wird, wenn er bis zum Gerichtstermin auf Kautions freikommt. In Studien konnten mittlerweile zwei Aspekte zum Einsatz des Algorithmus beleuchtet werden: Der Algorithmus benachteiligt systematisch schwarze Bürger, obwohl die Hautfarbe nicht Teil des Fragenkatalog ist. Und im direkten Vergleich mit einem einfachen Algorithmus, der lediglich fünf Datenpunkte zur Verfügung hatte, konnte der KI-Algorithmus keine besseren Angaben zur Rückfallwahrscheinlichkeit ermitteln.³

Vorausschauende neuronale Netzwerke

Eine weitere Vorhersagemethode ist die Verwendung Künstlicher Neuronaler Netzwerke. Aktuell werden für Wettervorhersagen neuronale Netze entwickelt, die einerseits die Daten von Wettermessstationen analysieren, aber auch lokale Messgrössen einbeziehen, die bei heutigen Prognoseprogrammen nicht berücksichtigt werden. Die Künstlichen Neuronalen Netzwerke sind in der Lage, auch die nicht linearen und chaotischen Wechselwirkungen von Wetterphänomenen in der Vorhersage zu verarbeiten. Bei lokalen Kurzzeitprognosen sind die KI-Systeme bereits heute den klassischen Modellen überlegen.⁴ Wie bei allen KI-Anwendungen ist auch bei den Prognosen die Qualität der Trainingsdaten entscheidend in Bezug auf die Qualität der Ergebnisse. Insbesondere bei Wettervorhersagen, wo grosse Datenmengen aus der Vergangenheit und der Gegenwart existieren, müssen die Anforderungen an Datensätze für das Training von Künstlichen Neuronalen Netzwerken hoch sein. Für Meteorologen spielt die Nachvollziehbarkeit der Prognosen etwa entlang von thermodynamischen Gleichungen eine zentrale Rolle. Es bleibt abzuwarten, wie die Ergebnisse von Künstlichen Neuronalen Netzwerken, deren Entstehung nicht transparent ist, sich innerhalb der Wettervorhersagen als fundiert etablieren.

Prognosen durch Fehler im Muster

Vorhersagen können mit KI-Systemen ferner getroffen werden, indem nach einer Mustererkennung neue Datenpunkte in

Bezug auf ihre Konsistenz mit dem Muster überprüft werden. In einer Applikation zur Früherkennung von Rückfällen von Drogenabhängigen wird die Nutzung des Smartphones zur Mustererkennung herangezogen: Aus Geodaten, E-Mail-Verkehr, Telefonaten und Internetkonsum werden Routinen abgeleitet, die auf eine stabile Situation der Patienten schliessen lassen. Ändert eine Person nun ihre Routinen (Telefonate zu ungewöhnlichen Zeiten, neue Routen usw.), benachrichtigt die Applikation die Betreuer und weist auf die erhöhte Rückfallgefahr hin. Mit 92% Genauigkeit soll das System die Rückfallgefahr der nächsten drei Tage bestimmen können.⁵

ANWENDUNGEN HEUTE UND MORGEN

- **Gesundheit:** Künstliche Intelligenz kann die Wahrscheinlichkeit für Herzinfarkte und Schlaganfälle in einem bestimmten Zeitrahmen anhand von Retinaanalysen bestimmen.
- **Industrie:** Durch Künstliche Intelligenz kann die prädiktive Analytik besser voraussagen, wie sich die Verkaufszahlen und Produktions- und Unterhaltskosten für ein Unternehmen entwickeln werden.
- **Mobilität:** Anhand von KI-basierten Vorhersagen können Verkehrsüberlastungen besser vorhergesagt und Verkehrsteilnehmer frühzeitig gelenkt werden.
- **Medien:** Algorithmen können Reichweiten von neuen Social Media Beiträgen vorhersagen.



KOMMUNIZIEREN

Conversational Interfaces sind automatisierte Computersysteme, mit denen Menschen schriftlich oder mündlich kommunizieren können. Die Kommunikationsfähigkeit intelligenter Systeme beruht auf dem Gebiet Natural Language Processing (zu Deutsch: Computerlinguistik). Die Disziplin befasst sich mit Methoden zur Beschreibung von Regeln für den menschlichen Informationsaustausch, die ein Computer anwenden kann. Kommunikative Assistenzsysteme haben in den letzten Jahren durch den Einsatz von Künstlicher Intelligenz in der Verarbeitung, Analyse und Generierung von Texten und gesprochener Sprache Fortschritte erzielt. Frühere Versuche, die Regeln für den menschlichen Informationsaustausch manuell zu definieren, scheiterten an der hohen Komplexität des menschlichen Sprachgebrauchs. Mehrdeutige Begriffe und Sarkasmus sind nur zwei der vielen Herausforderungen für das Erlernen einer natürlichen Sprache. Erst durch den Einsatz von Machine-Learning-Methoden wie Künstlichen Neuronalen Netzwerken konnten KI-Systeme zahl-

reiche text- und audiobasierte Inputs sinnvoll interpretieren, wobei die Nuancen des menschlichen Sprachgebrauchs immer noch Limitationen darstellen.

Chatbots

Textbasierte Dialogsysteme, sogenannte Chatbots, simulieren natürliche Unterhaltungen mit Menschen und werden beispielsweise in der Kundenbetreuung eingesetzt. Der Einführung solcher digitalen Assistenten liegt die Annahme zugrunde, dass sich ein Grossteil der Anliegen von Privat- und Geschäftskunden oder auch Patienten im Gesundheitsbereich automatisiert beantworten lässt. Einer der wesentlichen Vorteile von digitalen Dialogsystemen besteht in ihrer ständigen Verfügbarkeit. Ein weiteres Potenzial liegt in der Effizienzsteigerung durch die Automatisierung der Beratungsaufgaben, da standardisierte und repetitive Anliegen nicht mehr länger von physischen Mitarbeitenden entgegengenommen werden müssen.

Aktuell sind Chatbots in der Lage, die am häufigsten gestellten Fragen zu Produkten oder Dienstleistungen dynamisch zu beantworten oder bei Bedarf die Anliegen zu den richtigen Anlaufstellen weiterzuleiten. Gesundheitsapps bieten beispielsweise eine textbasierte Symptomüberprüfung als Frage-Antwort-Unterhaltung an und stellen je nach Zustand des Patienten relevante Informationen zur Verfügung. Falls der Chatbot das Anliegen des Nutzers nur unzureichend bearbeiten kann, empfiehlt das System einen Arztbesuch.⁶

Patienten- und Kundenbetreuung sind jedoch nicht die einzigen Anwendungsbereiche, in denen textbasierte Dialogsysteme zum Einsatz kommen. Aktuell werden vermehrt Lösungen entwickelt, die verschiedene Arten von Konversationen simulieren. Dabei können sich Nutzer durch die Interaktion mit Chatbots auf Bewerbungsgespräche, Geschäftsverhandlungen oder im Privatleben auf das Beenden einer Beziehung vorbereiten. Beim Einsatz von Chatbots für die Behandlung von Depressionen konnte in Pilotstudien belegt werden, dass vor allem bei Erstkonsultationen die Patienten gegenüber dem Chatbot offener und ehrlicher antworten als gegenüber einem Arzt. Die Wirksamkeit einer längerfristigen Psychotherapie entsteht jedoch über den Aufbau von Vertrauen zwischen Patient und Arzt, was auch mittelfristig nicht über einen Chatbot simuliert werden kann.⁷

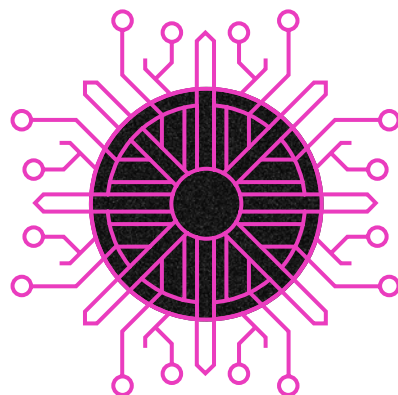
Persönliche Sprachassistenten

Im Gegensatz zu Chatbots, bei denen die Bedienung einer Tastatur oder eines Bildschirms notwendig ist, können Sprachassistenten gesprochene Kommunikation zwischen Menschen und Maschinen ermöglichen. Sie interpretieren gesprochene Inputs und generieren mittels synthetischer Stimmen Antworten auf die Anliegen der Nutzer. Apple lancierte mit Siri einen der ersten modernen Sprachassistenten im Jahr 2011, nachdem die gleichnamige Firma ein Jahr zuvor aufgekauft worden war.⁸ Seitdem kamen Sprachassistenten nicht mehr nur als Teil der Betriebssysteme von Smartphones, sondern vermehrt auch als eigenständige Haushaltsgeräte wie etwa

als Smart Speaker auf den Markt. Sie erlauben es unter anderem, Haushaltsinstallationen zu kontrollieren, Medien abzuspielen, den Tag mittels Kalender, Wetterdienst und Routenplaner zu organisieren oder Internetwissen abzurufen. Zudem werden spezialisierte Sprachassistenten entwickelt, die beispielsweise im Auftrag von rekrutierenden Unternehmen geeignete Kandidaten anrufen, Termine vereinbaren und anhand von Stellenbeschreibungen erste Interviews führen.⁹

ANWENDUNGEN HEUTE UND MORGEN

- **Tourismus:** Chatbots helfen in Museen und Ausstellungen Sachverhalte zu erklären, beantworten Fragen und navigieren durch Räume.
- **Versicherung:** Textbasierte Dialogsysteme übernehmen die Abwicklung von Schadensmeldungen.
- **Gesundheit:** Sprachassistenten mit VR-Avataren dienen als Gesprächspartner.
- **Retail:** Kommunikative Assistenzsysteme beantworten Fragen des Alltags und ermöglichen die sprachgesteuerte Bedienung von Haushaltsinstallationen.
- **Bildung:** Dialogsysteme ersetzen teilweise menschliche Lehrkräfte, gestalten Wissensvermittlung interaktiv und personalisieren Lerninhalte.
- **Gesundheit:** Die nächste Generation von Sprachassistenten hat das Potenzial, durch authentische Konversationen dezentrale Therapien zu simulieren.



BEWEGUNGEN STEuern

Künstliche Intelligenz kann den Menschen in der physischen Welt unterstützen, indem sie die mechanischen Bewegungen von Robotern in Umgebungen mit sich ändernden Rahmenbedingungen lernfähig und flexibel steuert. Dies ist nicht nur bei autonomen Fahrzeugen im Strassenverkehr, sondern auch bei Drohnen in Kriegsgebieten oder intelligenten Produktionsrobotern in der Automobilindustrie relevant. Durch Sensoren können Roboter analoge Informationen in digitale Daten umwandeln und aufgrund von Mustererkennung interpretieren. Audiosignale können ebenso wie visuelle Inputs oder Temperaturangaben dazu verwendet werden, um einem Roboter die Verarbeitung seiner Umgebung zu ermöglichen. Künstliche Intelligenz eignet sich zur Steuerung von verschiedenen Arten von Robotern. Die meisten Praxisbeispiele lassen sich einer von drei grundlegenden Kategorien zuordnen.¹⁰

Statische Manipulatoren

Manipulatoren umfassen statische Roboter, die ohne Navigationsfähigkeit an einem bestimmten Ort physische Handlungen vornehmen. Insbesondere bei der Automatisierung von ortsgebundenen Arbeitsschritten kommen Manipulatoren zum Einsatz. Ein Beispiel sind intelligente Industrieroboter, die in einer Fertigungsanlage durch Sensoren Daten aufnehmen und verarbeiten, aus ihrer Erfahrung lernen und autonom Entscheidungen fällen. In der Automobilindustrie können intelligente Manipulatoren Ungenauigkeiten in der Zusammensetzung von einzelnen Komponenten feststellen und diese ohne menschliche Intervention korrigieren. In Krankenhäusern können Operationsroboter Eingriffe ausführen, die die feinmotorischen Fähigkeiten von erfahrenen Chirurgen übersteigen.¹¹

Mobile Roboter

In Abgrenzung zu den Manipulatoren sind mobile Roboter in der Lage, sich fortzubewegen. Ob Wasser, Land, Luftraum oder Weltall: Die Verwendung von Künstlicher Intelligenz unterstützt insbesondere die Navigationsfähigkeit von autonomen Vehikeln. Ein prominentes Beispiel sind selbstfahrende Fahrzeuge im Strassenverkehr, bei denen ein komplexes Zusammenspiel von Kameras mit Bilderkennungssoftware, Lasersensoren und GPS heute zumindest unter Laborbedingungen eine unbemannte Fortbewegung ermöglichen. Seit 2012 erkundet der Marsroboter Curiosity den roten Planeten und ermittelt autonom, gestützt auf selbstlernender Software, geeignete Gesteinsproben. Dabei wird der

«In Zukunft werden wir ganz neue Arten von Maschinen und Robotern sehen, die uns im Alltag begleiten und mit uns arbeiten.»

Jana Koehler, Informatikprofessorin und wiss.-techn. Geschäftsführerin des Deutschen Forschungszentrums für KI

fast eine Tonne wiegende Forschungsroboter zwar von Menschen navigiert, die Künstliche Intelligenz kommt jedoch bei der Selektion und Steuerung des Fernerkundungsinstruments zum Einsatz.¹²

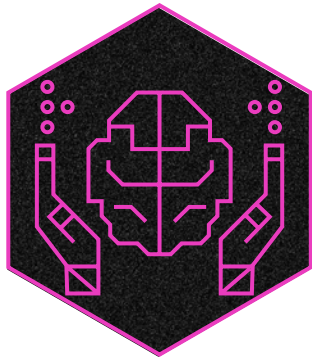
Humanoide Maschinen

Humanoide Roboter orientieren sich an der Physiologie des menschlichen Körpers. Typischerweise ist die Gestalt den Beinen, Oberkörpern, Armen und Köpfen von Menschen nachempfunden. Neben der mechanischen Konstruktion sollen die motorischen Fähigkeiten und Bewegungsabläufe des Menschen möglichst realitätsgetreu abgebildet werden. Künstliche Intelligenz ermöglicht humanoiden Robotern das Erlernen eines breiten Spektrums an Bewegungen: laufen, rennen, fangen, werfen, transportieren und gestikulieren beschreiben eine Auswahl an motorischen Fähigkeiten. Darüber hinaus werden die Maschi-

nen vermehrt auch mit kommunikativen Assistenzsystemen ausgestattet, um menschenähnliche und dialogfähige Androide zu schaffen. Letztlich kann Künstliche Intelligenz auch zur Steuerung von bionischen Prothesen verwendet werden. Dabei ermöglichen Machine-Learning-Algorithmen beispielsweise die beabsichtigten Handlungen eines Menschen innerhalb von Millisekunden von Muskelkontraktionen in konkrete Fingerbewegungen von Handprothesen zu übersetzen.¹³

ANWENDUNGEN HEUTE UND MORGEN

- **Gesundheit:** Pflegeroboter unterstützen Physiotherapeuten bei der manuellen Therapie und passen ihre Bewegungen den jeweiligen Patienten an.
- **Sicherheit:** Patrouillen bestehend aus autonomen Drohnen und Fahrzeugen überwachen Areale und Gebäude.
- **Sport:** Roboterisierte Trainingsgeräte bieten direktes Feedback und Lernassistenz.
- **Logistik:** Autonome Roboter liefern E-Commerce-Bestellungen in ländlichen Regionen aus.
- **Landwirtschaft:** Ernteroboter ermitteln durch Bildanalyse den Reifegrad.



GENERIEREN

Wenn Künstliche Intelligenz etwas Neues erschafft, kommt sie den menschlichen Vorstellungen von Kreativität und Innovation am nächsten. Neue Problemlösungsansätze oder Kreationen können durch ein KI-System beispielsweise dann generiert werden, wenn sie in einem Datensatz ein bestimmtes Muster erkennen und anhand der gelernten Regeln eine neue Variation durch die wiederholte Anwendung des Musters erschaffen. Der bisherige Einsatz der generierenden Künstlichen Intelligenz erstreckt sich nebst der bildenden Kunst über weitere Sektoren von der Raumplanung, der Herstellung neuer Geschmacksnoten bis hin zur Entwicklung neuer Sprachen und Therapieformen.

Kunstwerke kreieren

Durch Mustererkennung von Notensätzen kann Künstliche Intelligenz die Musikharmonien eines bestimmten Komponisten entschlüsseln und mit den gelernten Regeln neue Musikstücke komponieren, die derselben musikalischen Färbung angehören.¹⁴ Ähnliche Anwendungen finden sich bei

den bildenden Künsten: Trainiert man einen Algorithmus mit einer Datensammlung von Bildern, können daraus neue Entwürfe von Gemälden entstehen, die dann maschinell gemalt bzw. schichtweise aufgetragen werden. Das von einer Künstlichen Intelligenz gezeichnete Porträt eines Manns im Stil Rembrandts wurde als so gut befunden, dass das Werk innerhalb der Rembrandt-Werke nicht als fremdartig und maschinell produziert zu erkennen war. Den porträtierten Mann hat es so jedoch nie gegeben.¹⁵

Konzepte vorschlagen

Nebst der Gestaltung von Kunstwerken kann Künstliche Intelligenz auch in anderen Sektoren für innovative und kreierende Tätigkeiten eingesetzt werden. Im Gesundheitsbereich können KI-Systeme durch Mustererkennung Korrelationen bei Therapieverläufen herausarbeiten und darauf basierend neue, individuelle Behandlungsvorschläge erarbeiten. Bei Videospielen werden KI-Anwendungen eingesetzt, die das individuelle Spielverhalten analysieren und daraus kontinuierlich und in Echtzeit neue Spielinhalte generieren. In der physischen Welt hat ein KI-Modell die neue Sportart Speedgate – eine Mischung aus Fussball und Rugby – erfunden. Das Künstliche Neuronale Netzwerk wurde mit Charakteristika von über 400 Sportarten trainiert und entwickelte darauf basierend neue Spielkonzepte samt Regelwerk.¹⁶ Beim generativen Design wird durch Designer, Entwickler oder Ingenieure ein gestalterisches Ziel entlang von Parametern wie zum Beispiel Objektgrösse,

Kostenziele und Gewicht definiert. Die KI-Software entwickelt unter Berücksichtigung der Parameter Entwurfsalternativen, die schliesslich von Menschen ausgewählt und prototypisiert werden können. Die Anwendungen reichen von leichteren Bauteilen in der Luft- und Raumfahrt bis zu architektonischen Konzepten zur maximalen Raumausnutzung.

Deep Fakes

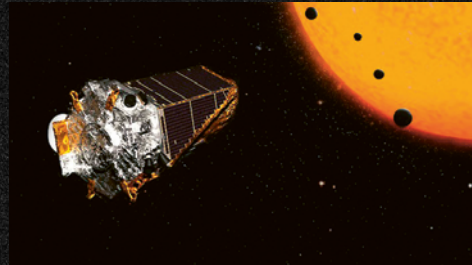
Ein spezifischer Anwendungsbereich von generativer Künstlicher Intelligenz ist die sogenannte Deep-Fake-Technologie, bei der Medieninhalte durch Künstliche Neuronale Netzwerke autonom erzeugt werden, um Bild-, Video- oder Audioaufnahmen zu fälschen bzw. zu manipulieren. Einige wenige reale Aufnahmen genügen oftmals als Basis, um innerhalb von wenigen Sekunden komplett neue Inhalte zu kreieren, die ohne systematische Analyse nicht mehr von Originalen unterscheidbar sind.¹⁷ Durch Inhalte, die mit Deep-Fake-Technologie produziert werden, können beispielsweise öffentliche Personen mit gefälschten Aussagen oder Handlungen diskreditiert werden, was insbesondere im Kontext von demokratischen und partizipativen Entscheidungsprozessen eine Herausforderung darstellt.

ANWENDUNGEN HEUTE UND MORGEN

- **Kunst:** Künstliche Neuronale Netzwerke können durch Mustererkennung in bestehenden Werken neue Varianten in der bildenden Kunst und der Musik erschaffen.
- **Lebensmittelindustrie:** Selbstlernende Algorithmen können neue Kombinationen und Formeln von Geschmacksnoten herstellen.
- **Gesundheit:** Künstliche Intelligenz identifiziert durch Mustererkennung Korrelationen bei Vitalwerten und legt damit die Basis für neue Therapieformen.
- **Unterhaltung:** Künstliche Neuronale Netzwerke generieren in Computerspielen anhand von Spielentscheidungen des Menschen neue Möglichkeiten für den Fortgang des Spiels.
- **Pharma:** Generative Chemie ermöglicht eine effizientere Entwicklung von Lead-Substanzen beispielsweise für neue Krebstherapien.
- **Kreativindustrie:** Ganzkörper-Deep-Fakes können für Spieldesign und Filmgestaltung verwendet werden.

FALLBEISPIELE AUS DER PRAXIS

Künstliche Intelligenz ist bereits heute im Einsatz bei Anwendungen im privaten und professionellen Umfeld. Die folgenden Fallbeispiele zeigen exemplarisch den aktuellen Entwicklungsgrad rund um den Praxiseinsatz der nächsten Generation von Computersystemen: Dazu zählen Chatbots als therapeutische Helfer, automatisierte Hautanalysen durch Bilderkennung oder die Generierung von Kunstwerken basierend auf Künstlichen Neuronalen Netzwerken.



1. Unbekannte Planeten identifizieren

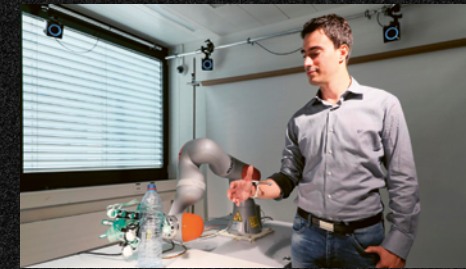
Mit dem Kepler-Weltraumteleskop der NASA wird durch die Transitmethode nicht der eigentliche Planet beobachtet, den es zu entdecken gilt, sondern die Sonne, die er umkreist. Schiebt sich ein Exoplanet nämlich vor seine Sonne, kommt auch weniger Licht beim Teleskop an. Der Algorithmus hilft, diese verringerte Helligkeit als Signal zu registrieren, um dadurch neue Planeten und ihre Eigenschaften zu identifizieren.

<https://www.vice.com/de/article/xwvavz/kunstliche-intelligenz-beweist-unser-sonnensystem-ist-nicht-einzigartig-im-all>

2. Chatbot als erster Ansprechpartner bei Angstattacken

Die App Replika vom amerikanischen Softwareunternehmen Luka stellt in einem Chat eine Künstliche Intelligenz als Kommunikationspartner bereit, mit dem man wie mit einem Freund chatten kann, der Fragen stellt und Fragen beantwortet. Das System lernt im Lauf der Unterhaltung und baut das neue Wissen in die Konversation ein, damit es ein immer besserer Gesprächspartner und wenn erwünscht auch ein therapeutischer Helfer in schwierigen Situationen werden kann.

<https://replika.ai/>



3. Roboterprothesen durch Muskelsignale steuern

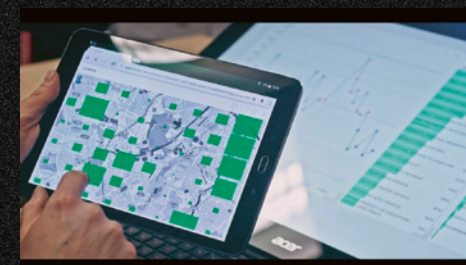
EPFL-Wissenschaftler entwickeln neue Ansätze zur verbesserten Steuerung von Roboterhänden für Personen mit amputierten Händen. Durch die Zusammenarbeit von Neuroengineering und Robotik gelingt es, die beabsichtigte Fingerbewegung aus der Muskelaktivität am stumpfen Arm zu entschlüsseln und so eine präzise Fingerbewegung der prothetischen Hand zu ermöglichen. Es kommt ein maschineller Lernalgorithmus zum Einsatz, der individuell an Patienten trainiert werden kann.

<https://actu.epfl.ch/news/a-smart-artificial-hand-for-amputees-merges-user-a/>

4. Hautanalyse mit Bilderkennungssoftware bewirbt Pflegeprodukte

Der Hautdiagnose-Chatbot S.A.L. von Ponds interpretiert durch Fotos mit der Smartphone-Kamera die Haut des Gesichtes. Mittels Mustererkennung wird die Hautgesundheit durch vier Indikatoren klassifiziert: Gleichmäßigkeit vom Hautton, Flecken, Pickel und Falten. Durch die Hautanalyse der Fotos soll der Chatbot über die Hautgesundheit aufklären und Pflegeempfehlungen abgeben.

<https://www.btbp.org/News.html>



5. Verbrecher erfolgreicher mit Künstlicher Intelligenz jagen

Für die Unterstützung der Arbeit von Kriminalpolizisten hat die Middlesex-Universität das System VALCRI entwickelt. In einer Datensammlung über Verhöre, Entscheide, Bilder und Videos wird mithilfe eines Algorithmus nach Ähnlichkeiten oder Verbindungen gesucht, die für einen aktuellen Fall relevant sein könnten. Gefundene Parallelen werden entsprechend visualisiert, um daraus Ermittlungsvorschläge abzuleiten.

<http://valcri.org/>



6. Vitalfunktionen von Frühgeborenen neu deuten

Die Universität Ontario hat zusammen mit IBM eine Software entwickelt, die Ärzten bei der medizinischen Versorgung von Frühgeborenen helfen soll. In Datensammlungen von Vitalfunktionen und Infektionsausbrüchen vergangener Fälle konnte der Algorithmus ein Muster identifizieren. Das entdeckte Muster ist für Ärzte kontraintuitiv: Denn die Vitalwerte eines Frühgeborenen vor einer Infektion sind überraschenderweise stabil. In der Anwendung kann das Muster mit den Echtzeitvitalwerten eines Frühgeborenen verglichen werden, um so eine lebensgefährliche Infektion vor dem Auftreten der Symptome zu erkennen.

Mayer-Schönberger, Viktor; Cukier, Kenneth (2017): *Big Data. Die Revolution, die unser Leben verändern wird*. 3. Auflage. München, S. 79. oder online: <https://www.ibmbigdatahub.com/sites/default/files/document/ODC03157USEN.PDF>

7. Ein Roboter zum Burgerbraten und Esseneinpacken

Der robotische Küchenassistent Flippy entstand aus der Zusammenarbeit zwischen CallBurger und Miso Robotics. Mit einem sechsachsigen Arm wendet er das Burgerfleisch auf einem Grill, gibt Beilagen hinzu, verpackt den Burger und informiert die Mitarbeitenden über die Fertigstellung.

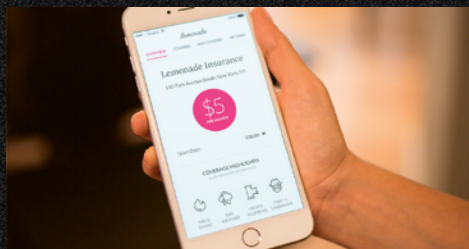
<https://www.manager-magazin.de/digitales/it/roboter-flippy-versagt-in-kueche-von-burger-restaurant-a-1198806.html>



8. Eine Versicherung mit Chatbot-Angestellten

Beim amerikanischen Versicherungsanbieter Lemonade sind mehrere interne Prozesse sowie der Austausch mit Kunden und Kundinnen durch Künstliche Intelligenz unterstützt. Während der Chatbot Maya die Bedürfnisse der Kunden und Kundinnen ermittelt, analysiert und ein Angebot erstellt, nimmt der Bot Jim Schäden auf, führt die Rückzahlung automatisch aus oder gibt komplexere Fälle an menschliche Kollegen weiter. Laut Lemonade dauert eine Schadenfallverarbeitung mit Rückzahlung eines Velodiebstahls nur drei Sekunden.

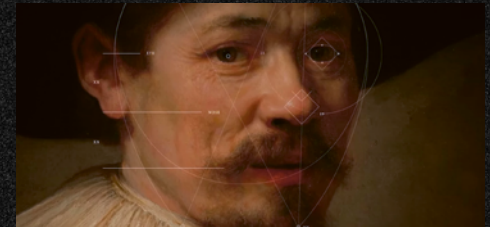
<https://www.lemonade.com/>



9. Mit Künstlicher Intelligenz das Werk eines Künstlers vergrößern

Das Projekt «The Next Rembrandt» der ING-Group und Microsoft hat mithilfe eines intelligenten Algorithmus ein Gemälde erstellt, das alle Merkmale eines originalen Rembrandt-Gemäldes hat. Der Algorithmus wurde mit 3D-Scans der Originalwerke Rembrandts trainiert und analysierte durch Mustererkennung die Kerneigenschaften der Bilder des Künstlers. In Kombination mit weiteren Regeln, die weniger den Stil, sondern die Muster der porträtierten Gesichter abbilden, entstand ein Porträt eines erfundenen Manns mit einem neuen Gesicht, das in einem Schichtverfahren gedruckt wurde.

<https://www.nextrembrandt.com/>



10. Mit dem Roboter durch unwegsames Gelände

Die ANYmal-Roboter von ANYbotics und der ETH Zürich stehen auf vier Carbonbeinen und bewegen sich trittsicher über schwieriges Gelände. Durch ein Künstliches Neuronales Netz werden in einer Simulation Tausende ANYmals gleichzeitig trainiert, was die Laufgeschwindigkeit, die Reaktionszeit und damit die Erholungszeit nach Stürzen der echten Roboter laufend verbessert. Die verschiedenen Modelle sind im Erste-Hilfe-Einsatz, in der Forst- und Landwirtschaft, nehmen an Roboterfußballspielen teil und führen in der Industrie Inspektionen durch.

<https://www.anybotics.com/>



11. Computerspiel entwickelt sich laufend neu

An der Northwestern-Universität wurde das Computerspiel GPT-Adventure entwickelt, das sich selbst während des Spiels je nach Spielverhalten weiterentwickelt. Als Grundlage dient ein Künstliches Neuronales Netz, das mit den Daten von anderen Text-Adventure-Spielen trainiert wurde. Der Spieler hat stets neue Situationen zu bewältigen und agiert durch einfache Textbefehle wie «open the door», die dann den Handlungsverlauf laufend neu bestimmen.

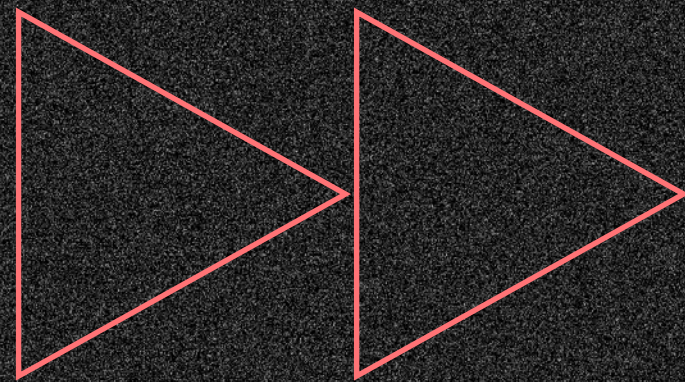
<https://quicktotheratcave.tumblr.com/post/187432425523/shall-we-play-a-game-a-gpt-2-text-adventure>

```
ride horse
THE HORSE LANDS GENTLY OF COURSE, EXCEPT FOR THE SPLASH OF
water.

get on horse and ride it
THE HORSE IS TOO WELL TRAINED TO GALLOP OFF THE MOAT
IT REARS BACK ON ITS HIND LEGS, WHINNYING IN CONFUSION

ride horse into the canyon
MOUNTAINS, LEADING SOUTHWEST AND SOUTHEAST
A ROAD LEADS SOUTHWARD FROM HERE

go south
```

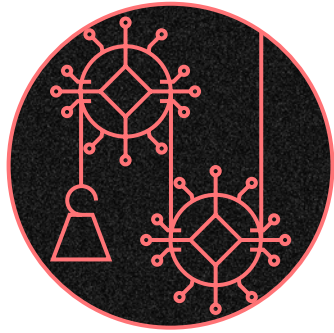



4

Konsequenzen

WAS KI FÜR DIE WELT VON MORGEN BEDEUTET

Der potenziell breite Einsatz von Künstlicher Intelligenz in wirtschafts- und gesellschaftsrelevanten Bereichen erfordert eine differenzierte Analyse der absehbaren Mehrwerte, aber auch der Schwierigkeiten. Auf der einen Seite stehen Erwartungen an eine Effizienzsteigerung von automatisierbaren Prozessen, mehr Sicherheit für den öffentlichen Raum und einen höheren Grad an Personalisierung für Konsumprodukte und Gesundheitsleistungen. Auf der anderen Seite führt die Implementierung künstlich intelligenter Prozesse – wie sie heute entwickelt werden – auch zu kritischen Entwicklungen, wie es die Intransparenz von algorithmischen Entscheidungen, deren Potenzial zur Diskriminierung und die fehlende Klärung von Haftungsfragen bereits heute andeuten. Für individuelle, politische und ökonomische Entscheidungen im Themenfeld von intelligenten Maschinen gilt es, beide Seiten der Folgen zu berücksichtigen, um gesellschaftliche Innovation nachhaltig zu fördern.



CHANCEN – WIE KI MENSCHEN UNTERSTÜTZEN KANN

Aufbauend auf der Analyse der technologischen Grundlagen sowie den unterschiedlichen Anwendungsbereichen der Künstlichen Intelligenz lassen sich sektorübergreifende Chancen für Wirtschaft, Wissenschaft, Politik und Gesellschaft ableiten. Eine der grundlegendsten Veränderungen betrifft den Einsatz von selbstlernenden Algorithmen und autonomen Computersystemen zur Produktivitätssteigerung und Verbesserung der Lebensqualität.

Gesteigerte Effizienz

Die KI-basierte Automatisierung von Aufgabenbereichen, die zuvor ausschliesslich von menschlichen Mitarbeitenden ausgeführt wurden, verspricht eine Steigerung der Effizienz bestehender Prozesse. Anders als bei vergangenen industriellen Revolutionen betreffen diese Umwälzungen Tätigkeiten aller Komplexitäts- und Bildungsstufen. Die Effizienzsteigerungen ergeben sich primär daraus, dass repetitive Tätigkeiten

an KI-Systeme ausgelagert und somit ökonomische Ressourcen wie Geld und Zeit gespart werden können. Die Realisierung dieses Potenzials ist jedoch komplex und setzt den Aufbau einer leistungsfähigen Infrastruktur und eine hochwertige Datengrundlage voraus. In der Übergangszeit sind grosse Investitionen nötig, um autonome Systeme zu entwickeln, zu trainieren, einzuführen und zu überwachen. Somit erscheinen die Szenarien, wonach KI-Systeme bereits kurz- bis mittelfristig die Mehrzahl menschlicher Arbeitskräfte ersetzen, als überzeichnet. Die Automatisierung von Tätigkeiten und Berufsgruppen, die intensiv mit Menschen arbeiten und dazu empathische Fähigkeiten benötigen, ist beispielsweise aus heutiger Perspektive nicht absehbar. Auch komplexe Aufgaben, die ein kontextuales Hintergrundwissen voraussetzen oder die auf interdisziplinären Problemlösungen beruhen, werden voraussichtlich mittel- bis langfristig mehrheitlich von Menschen ausgeführt. Darüber hinaus zeigen auch wissenschaftliche Untersuchungen, dass die Zusammenarbeit zwischen Mensch und Maschine die effizienteste Kollaborationsform darstellt.¹

Verbesserte Qualität

Durch die Reduktion von menschlichen Fehlerquellen und erhöhte Präzision kann die Qualität von Produkten oder Dienstleistungen durch den Einsatz Künstlicher Intelligenz verbessert werden. Dies ist nicht nur in der Produktion relevant, wo KI-basierte Systeme zur stringenteren und prädiktiven Qualitätskontrolle beitragen, sondern auch bei Operationsrobotern, die feinmotorische

Tätigkeiten bei komplexen medizinischen Eingriffen ausführen. Die prognostizierten Qualitätsverbesserungen basieren auf der Annahme, dass Künstliche Intelligenz das kognitive Leistungsniveau des Menschen künftig in eng begrenzten Problemfeldern übertreffen wird. Ein weiterer Vorteil ergibt sich aus der ständigen Verfügbarkeit von KI-basierten Dienstleistungen. Gerade digitale Leistungserbringungen wie die Diagnose von Hautkrankheiten durch mobile Bilderkennungsprogramme oder die kommunikativen Assistenzsysteme in der Kundenbetreuung führen zu einem verbesserten Zugang, der orts- und zeitunabhängig gewährleistet werden kann.

Verstärkte Personalisierung

Durch die Berücksichtigung von bewussten und unbewussten Präferenzen sowie individuellen Merkmalen können KI-Systeme einen erhöhten Grad der Personalisierung erreichen. Empfehlungssysteme, die relevante Inhalte oder Produkte vorschlagen, sind wie im Fall von Netflix oder Amazon bereits fester Bestandteil aktueller Geschäftsmodelle. Durch die Automatisierung dieser Vorgänge ist es möglich, massgeschneiderte Produkte und Dienstleistungen in den Massenmarkt zu bringen. Gerade der erfolgreiche Einsatz von personalisierten Empfehlungssystemen mit integrierter Lernfähigkeit stellte in den letzten Jahren einen entscheidenden Wettbewerbsvorteil dar. Laut Schätzungen erwirtschaftete der Grosskonzern Amazon ein Drittel seines Gesamtumsatzes aufgrund von personalisierten Produktempfehlungen.² Künftig könnte die KI-basierte

Personalisierung auch vermehrt im Gesundheitsbereich Verbreitung finden, wenn präventive Massnahmen und medizinische Behandlungen auf die individuellen Merkmale der Patienten abgestimmt werden.

Erhöhte Sicherheit

Eine weitere Chance besteht in der Gewährleistung und Erhöhung der menschlichen Sicherheit. Durch die Vermeidung von gefährlichen Tätigkeiten wie dem Aufspüren und Entschärfen von Minen in Kriegsgeländen, bei denen sich Menschen hohen Risiken aussetzen, können physische Schäden verhindert werden. Autonome Robotersysteme erlauben es, solche Aktivitäten vermehrt ohne menschliche Beteiligung auszuführen. Im Gegensatz zu Menschen, die beispielsweise durch Ablenkungen oder unter Einfluss von Betäubungsmitteln Fehlentscheidungen treffen, haben KI-basierte Systeme zumindest theoretisch den Vorteil, sich konsequent an Vorgaben halten zu können. Im Kontext des autonomen Fahrens gehen die optimistischsten Schätzungen von einer Reduktion der Unfalltoten von bis zu 90% aus.³ Noch sind jedoch zahlreiche technologische und regulatorische Herausforderungen zu bewältigen, bevor die Zahl der Unfälle durch den Einsatz von selbstfahrenden Fahrzeugen signifikant reduziert werden kann.

Nicht nur in der physischen Welt kann die Sicherheit der Bevölkerung durch den Einsatz von Künstlicher Intelligenz erhöht werden. Auch im virtuellen Raum wie beispielsweise in den sozialen Medien können Videoerkennungssysteme Gewaltinhalte



identifizieren und automatisiert entfernen. Dies führt zu einer Entlastung von menschlichen Mitarbeitenden, die Medieninhalte nicht mehr länger auf Verstösse gegen die internen Richtlinien überprüfen müssen. Darüber hinaus kommen Machine-Learning-Algorithmen bei der Detektion von Anomalien zum Einsatz, um Betrugsfälle oder Cyberattacken aufzudecken.

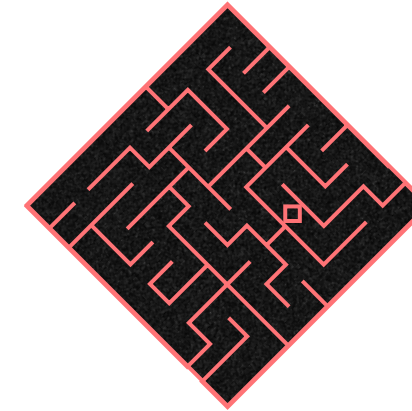
Erweiterte Kreativität

Künstliche Intelligenz kann einen Beitrag zur Innovationstätigkeit von Menschen leisten. Indirekt schafft Künstliche Intelligenz Raum für menschliche Innovationen, indem repetitive Aufgaben vermehrt ausgelagert werden und sich Menschen auf kreativere Tätigkeiten fokussieren können. Auf der anderen Seite können KI-Systeme neue Problemlösungsvorschläge entdecken, die dem Menschen bis anhin verwehrt blieben. So entwickelte Deep Minds mit dem Programm AlphaGo neuartige und unkonventionelle Taktiken, die auch von Grossmeistern des strategischen Brettspiels über Jahrzehnte ignoriert wurden.⁴ Auch in anderen Bereichen wie der Pharmazeutik, Raumplanung oder Verkehrsführung können KI-Systeme unter Einhaltung von klaren Vorgaben innovative Problemlösungsvorschläge erarbeiten, die die menschliche Ideenfindung und Vorstellungskraft ergänzen.

Geförderte Produktivität

Durch die KI-basierte Innovation von Produkten und Dienstleistungen kann in Kombination mit der Automatisierung von standardisierten Tätigkeiten die Produkti-

vität ganzer Wirtschaftszweige gesteigert werden. So geht die britische Unternehmensberatungsfirma PricewaterhouseCoopers nach eigenen Schätzungen davon aus, dass bis ins Jahr 2030 die globalen Produktivitätsgewinne durch Künstliche Intelligenz bis zu 15,7 Milliarden US-Dollar betragen werden.⁵ Auch wenn eine zuverlässige Schätzung der Wertschöpfung durch den Einsatz von KI-Systemen aus heutiger Perspektive schwierig erscheint, stehen solche Prognosen stellvertretend für die grosse Erwartungshaltung.



HERAUSFORDERUNGEN – WO KI DIE GESELLSCHAFT VOR PROBLEME STELLT

Aufgrund der Entscheidungsfähigkeit von KI-Systemen und der Auswirkungen, die sich daraus für den Menschen und seine Umwelt ergeben, sind ethische Fragestellungen von zentraler Bedeutung. Doch nach welchen ethischen Prinzipien sollten sich KI-Systeme entscheiden? Eine von der Europäischen Union einberufene Expertengruppe hat zur Beantwortung dieser Frage die folgenden vier Grundsätze vorgeschlagen: Respekt vor menschlicher Autonomie, Nichtschaden, Gerechtigkeit und Erklärbarkeit.⁶ Problematisch ist jedoch, dass die vier Grundsätze bei spezifischen Fragestellungen miteinander in Konflikt stehen können. Wie sollte ein Pflegeroboter entscheiden, der die Autonomie eines demenzenden Patienten beschneiden müsste, um ihm seine überlebenswichtigen Medikamente zu verabreichen?

Diskriminierende Entscheidungen

Die Diskriminierung von Personen oder Personengruppen bei KI-unterstützten Entscheidungsprozessen repräsentiert eines der zentralen Problemfelder. Die Ungleichbehandlung von Individuen ohne nachvollziehbare Begründung kann beispielsweise daher kommen, dass Korrelationen, die nicht mit einem Kausalzusammenhang in Verbindung stehen, fälschlicherweise bei KI-basierten Entscheidungen berücksichtigt werden. Um die Ergebnisse statistisch zu optimieren, können sich KI-Systeme zudem unter Umständen zu stark an die Trainingsdatensätze anpassen. So musste Amazon seine interne Rekrutierungssoftware einstellen, weil diese aufgrund der historischen Trainingsdaten Frauen diskriminierte.⁷ Neben der Replikation von historischen Verzerrungen kann auch die fehlende Diversität im Entwicklerteam einen Grund für Diskriminierungstendenzen darstellen, da spezifische Geschlechter, Ethnien oder Altersgruppen in der Entwicklung von KI-Systemen über- bzw. unterrepräsentiert sind. Um die Akzeptanz von KI-basierten Anwendungen zu steigern, ist es deshalb zentral, diskriminierende Tendenzen früh zu erkennen und sie durch technologische Weiterentwicklungen und strenge Auflagen zu minimieren.

Intransparente Systeme

KI-Anwendungen zur Entscheidungsfindung sind mit einer doppelten Intransparenz konfrontiert. Einerseits ist es für Nutzer und Nutzerinnen nicht ersichtlich, ob bzw. wann sie mit einem KI-System in Kontakt kommen. Dies gilt für KI-basierte Rekrutierungsentscheidungen, automatisierte Kreditver-





gaben oder Chatbots zur Kundenbetreuung. Die zweite Ebene fehlender Transparenz betrifft die Funktionsweise der Technologie. Die Erklärbarkeit von KI-Systemen ist nicht zwangsläufig gegeben, da Machine-Learning-Verfahren eine Vielzahl von Faktoren berücksichtigen und deren Ergebnisse nicht immer nachvollziehbar dargelegt werden können. Die Forschung kann massgeblich dazu beitragen, die Verunsicherung in der Bevölkerung im Umgang mit KI-Systemen einzudämmen. Aber nicht nur die Algorithmen selbst können intransparent sein. Für Unternehmen kann die Funktionsweise der Software ein Differenzierungsmerkmal und geistiges Eigentum darstellen. Es wird zu verhandeln sein, inwieweit Anbieter einer Forderung nach Transparenz zu den Mechanismen hinter Produktempfehlungen und Therapievorschlägen nachkommen werden müssen.

Erhöhte Verunsicherung

Auch der Einsatz von kommunikativen Assistenzsystemen oder KI-gesteuerten Robotern kann zu Verunsicherung innerhalb der Bevölkerung führen. So hat der japanische Roboteriker Masahiro Mori bereits in den 1970er-Jahren ein Phänomen beschrieben, wonach die Akzeptanz von Robotern und anderen technischen Objekten mit deren Menschenähnlichkeit zuerst stark ansteigt. Dieser Effekt besteht jedoch nur bis zu einem gewissen Grad. Sobald ein Roboter oder andere technische Objekte das menschliche Verhalten zu akkurat simulieren, sinkt ihre Vertrauenswürdigkeit abrupt, was neben Verunsicherung auch grosse Ängste auslösen kann.⁸ Darüber hinaus verbreitet sich mit dem tech-

nologischen Fortschritt auch die Grundskepsis, dass man sich nicht mehr auf die Authentizität von bestimmten Inhalten verlassen kann. Historisch waren Bewegtbild- und Audioaufnahmen sehr schwierig zu manipulieren. Mit sogenannten Deep Fakes sinken jedoch die Hürden, um gefälschte Videos oder Telefonate zu erzeugen und damit ausgewählte Zielgruppen zu erreichen und zu manipulieren. Die missbräuchliche Verwendung von KI-Technologien, um politische oder wirtschaftliche Vorteile zu erlangen, wird dadurch zu einer zentralen Herausforderung.

Unklare Verantwortlichkeiten

Eine weitere Herausforderung besteht darin, dass die Verantwortlichkeit und Haftung beim Einsatz von KI-Anwendungen noch weitestgehend unklar sind. Haften Ärzte, wenn sie sich auf ein medizinisches Diagnosesystem verlassen oder doch der Hersteller des KI-Systems? Gerichte haben in der Vergangenheit oftmals zugunsten der Ärzte entschieden und den Anwendungen eine ähnliche Rolle wie medizinischer Fachliteratur beigemessen.⁹ Dies könnte sich jedoch ändern, sobald medizinische Expertensysteme konstant das menschliche Leistungsniveau übersteigen. So argumentierte der US-Forscher Atul Gawande bereits um die Jahrtausendwende, dass menschliche Spezialisten zur Rechenschaft gezogen werden könnten, falls sie die Hilfe von deutlich überlegenen KI-Systemen nicht in Anspruch nehmen sollten.¹⁰ Aus heutiger Perspektive ist die ungeklärte Frage nach der Haftbarkeit ein wesentlicher Grund, weshalb der Einsatz von Künstlicher Intelligenz einen zusätzlichen menschlichen Entscheidungs-

schrift erfordert. Darüber hinaus führt die ungeklärte Lage im Kontext der Haftbarkeit beim Einsatz von KI-Systemen dazu, dass die Rechtssicherheit für Unternehmen und Institutionen heute nicht immer gewährleistet werden kann.

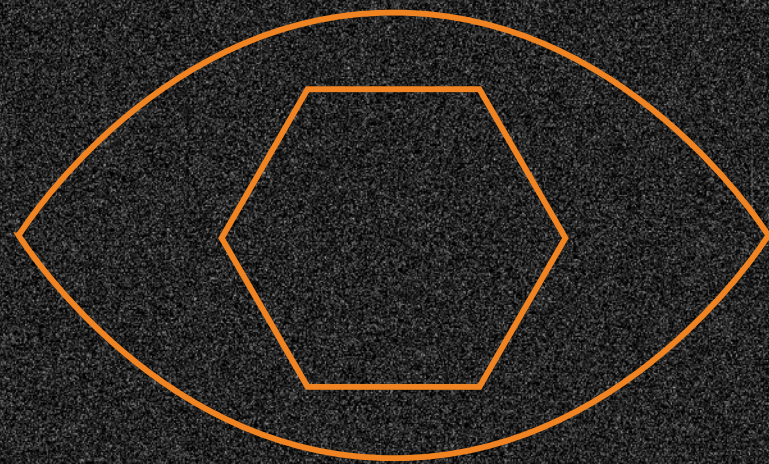
Gefährdete Privatsphäre

Durch den Einsatz grosser Datenmengen zur Verbesserung von KI-Lernmodellen steigt auch der Anreiz zur Anhäufung von sensiblen Daten bei Unternehmen und öffentlichen Institutionen. Der Schutz der Privatsphäre beschreibt in diesem Kontext eine fundamentale Herausforderung, da beispielsweise KI-basierte Vorhersagen von Krankheitsrisiken oder Berechnungen individueller Versicherungspolicen durch selbstlernende Algorithmen auf persönliche Daten angewiesen sind. Ein gewisses Konfliktpotenzial entsteht insbesondere dadurch, dass viele Chancen der Künstlichen Intelligenz gerade eben erst durch den Einsatz persönlicher Daten realisiert werden können. Deshalb wird die informationelle Selbstbestimmung im Kontext der Erhebung, Speicherung, Verwendung und Weitergabe von Nutzerdaten in Zukunft an Relevanz gewinnen. Dabei spielen auch Konzepte wie das Recht auf Unwissenheit eine Rolle, bei dem Personen vor Informationen geschützt werden, die die individuell gewünschte Lebensführung negativ beeinträchtigen können – etwa im Fall einer statistisch erhöhten Gefahr für eine Alzheimererkrankung, die sich aus der Analyse individueller Daten ableiten lässt. Die Regulierungen zum Schutz der Privatsphäre sind noch nicht auf das Zeitalter der Künstlichen Intelligenz abge-

stimmt. Dabei entstehen im Hinblick auf die informationelle Selbstbestimmung kontinuierlich neue Herausforderungen, wie beispielsweise die Schwierigkeit KI-Systemen beizubringen, verwendete Daten wieder aus einem Lernmodell zu löschen.¹¹

Homogene und extreme Outputs

Bei der Personalisierung von Inhalten auf sozialen Medien oder bei Suchmaschinen werden die Ergebnisse basierend auf verganginem Verhalten ausgewählt. Dies führt einerseits zu sogenannten Filterblasen, indem Nutzern und Nutzerinnen Informationen angezeigt werden, die ihren bisherigen Ansichten entsprechen bzw. von ähnlichen Nutzerprofilen geteilt werden. Dadurch reduziert sich die Vielfalt von Inhalten und es sinkt die Wahrscheinlichkeit, mit abweichenden Meinungen konfrontiert zu werden, die die eigene Position infrage stellen könnten. Dasselbe gilt für Empfehlungen auf Musik- oder Videostreamingseiten, wo unabhängig vom Anbieter gestützt auf vergangenheitsbasierten Datenpunkten gleiche Empfehlungen gemacht werden. Im Fall des Videoportals Youtube konnte zudem nachgewiesen werden, dass KI-Algorithmen kontinuierlich extremere Videos auswählen, um die Verweildauer zu steigern, was die Verbreitung von Verschwörungstheorien und expliziten Inhalten fördert. Infolge der Debatte hat sich das Unternehmen dazu bekannt, proaktiver gegen die Verbreitung verbotener Inhalte vorzugehen.¹² Ob Videoinhalte, die gegen die Richtlinien verstossen, automatisiert erfolgreich identifiziert und entfernt werden können, ist heute noch nicht gesichert.



5 Thesen

ZUR ZUKUNFT INTELLIGENTER MASCHINEN

Basierend auf den sichtbaren und potenziellen Anwendungsfeldern der Künstlichen Intelligenz samt den weiterführenden Chancen und Herausforderungen gilt es die Zukunft der lernenden Maschinen in einer grösseren Perspektive weiterzudenken. Denn trotz der verschiedenen Limitationen und der Tatsache, dass die Erwartungen an KI-Systeme medial überbewertet sind, werden lernende Maschinen und smarte Algorithmen Teil der neuen Normalität. Allerdings werden die Einsatzgebiete vielmehr durch ihre Rolle als flexible und personalisierte Assistenzsysteme geprägt sein als durch Massen von humanoiden Robotern, die den Menschen ersetzen. Deshalb ist das zentrale Innovationsfeld nicht allein die Weiterentwicklung von Algorithmen und Hardware, sondern die Gestaltung von Schnittstellen zwischen Menschen und Maschinen. Es stehen auch gesellschaftliche Fragen im Zentrum, wenn etwa durch die Verbreitung von Deep-Fake-Videos das Vertrauen in die virtuelle Welt schrumpft oder wenn KI-Systeme durch tendenziöse Trainingsdaten zu Diskriminierung neigen. Gleichzeitig gilt es, die Möglichkeiten für wirtschaftliches Wachstum durch Effizienzsteigerungen oder neue Möglichkeiten in Bildung und Forschung zu nutzen. Denn wenn die Dynamik in Wirtschaft, Wissenschaft und Politik in den nächsten Jahren weiterhin hoch bleibt, braucht es dynamische und flexible Systeme, die in der Lage sind, in komplexen Umgebungen zu agieren. Dies ist der Mehrwert von Künstlicher Intelligenz – in Verknüpfung mit der menschlichen.



TOASTERS, NOT TERMINATORS

Intelligente Maschinen
werden zur Normalität, ihre
Kompetenzen sind jedoch limitiert.

Anwendungen der Künstlichen Intelligenz sind bereits seit Längerem im Alltags- und Berufsleben angekommen. Von der automatischen Sortierung von E-Mails oder der Verknüpfung von Menschen mit ähnlichen Interessen in sozialen Medien bis zu Musikempfehlungen bei Streamingdiensten sind Menschen täglich mit Lösungen von lernenden Maschinen konfrontiert. Die Zahl der Anwendungen wird in den nächsten Jahren weiter zunehmen und mehr Bereiche des privaten und beruflichen Lebens prägen. Dabei unterscheiden sich die effektive Form und Anwendung allerdings von den Bildern und Vorstellungen, die traditionell mit Künstlicher Intelligenz verknüpft sind. KI-Systeme sind keine Kopie humanoider Intelligenz, die Menschen in Beruf und anderen Rollen flächendeckend ersetzen. Es sind Unterstützungs- oder Cockpitsysteme, die eigenständig Aufgaben übernehmen und basierend auf Erfahrungen selbst Lösungen entwickeln können. Pointiert formuliert: Die Zukunft der Künstlichen

Intelligenz sind intelligente Toaster – und nicht Terminator-Roboter. Ihre Leistungsfähigkeit stösst bei der Lösung von Herausforderungen an Grenzen, die durch mehrdimensionale Herausforderungen, nicht quantifizierbare Prozesse oder fehlende, statistisch saubere und umfassende Datengrundlagen geprägt sind. Mit Blick auf die Einsatzpotenziale in der realen Welt spielt auch nicht allein die theoretische technische Leistungsfähigkeit eine Rolle, sondern die betriebswirtschaftliche Rentabilität. Um KI-Systeme nachhaltig einsetzen zu können, muss den Kosten ein entsprechender kommerzialisierbarer oder gesellschaftlicher Nutzen gegenüberstehen. Mit Blick auf die ferne Zukunft steht offen, wie weit Rechensysteme mit wesentlich höherer Leistungsfähigkeit – wie etwa Quantencomputer – eine neue Grundlage für eine Künstliche Intelligenz schaffen, die nicht mehr mit den heutigen Limitationen konfrontiert ist. Was fern in diesem Kontext heisst, lässt sich heute jedoch nicht seriös beantworten.





MACHINE TEACHING, NOT LEARNING

Künstliche Intelligenz erfordert Interaktion mit Menschen und Umgebung.

Der Einsatz Künstlicher Intelligenz ermöglicht verbesserte Prognosen, Effizienzsteigerungen und Optimierungsverfahren. Sie ist in der Lage, neue Wirkstoffe zu entwickeln oder sich in anspruchsvollen Strategiespielen durchzusetzen. Eine Künstliche Intelligenz grenzt sich von einfachen Algorithmen ab, indem sie in der Lage ist, nicht nur programmierte Prozesse anzuwenden, sondern diese – basierend auf Daten und Erfahrungen – zu optimieren und im übertragenen Sinn zu erlernen. Folglich entfalten KI-Systeme ihre Potenziale jedoch nur aufbauend auf dem bestehenden Wissensschatz der Umgebung oder in Wechselwirkung mit Menschen. Mit einem Perspektivenwechsel ist so gesehen nicht von lernenden Maschinen auszugehen, die selbstmotiviert neue Lösungen entwickeln, sondern vielmehr von einem Prozess des Machine Teaching, bei dem Menschen den Maschinen beibringen, wie Herausforderungen gemeistert werden können. Die wissenschaftliche Theorie des Embodiment geht ferner sogar davon aus, dass Maschinen effizienter lernen, wenn sie mithilfe eines eigenen Körpers realen Umgebungen

ausgesetzt sind: Ein Roboter, der über eine Türschwelle stolpert und daraus lernt, entwickelt sich entsprechend schneller als ein rein softwarebasiertes System. Darüber hinaus wurde bei der Entwicklung von Künstlichen Neuronalen Netzwerken gezeigt, dass die Systeme an offenen Herausforderungen besser lernen als bei spezifischen geschlossenen Fragestellungen. Mit dieser Perspektive ergeben sich neue Grundlagen für die Weiterentwicklung von Künstlicher Intelligenz, die einen zentralen gesellschaftlichen Aspekt einschließt: Interaktion mit dem Alltag ist essenziell. Einerseits um die effektive Leistungsfähigkeit für eine kontextuelle Wissensaufbereitung zu optimieren, andererseits als Grundlage für eine aktive und bewusste Sensibilisierung der Bevölkerung für KI-Systeme, die bislang mehrheitlich im Hintergrund zum Einsatz kamen. Unabhängig davon, wie autonom Systeme tatsächlich agieren, können die erhofften Potenziale nur dann eingelöst werden, wenn die Systeme in ihrem Zusammenspiel mit Menschen gedacht, entwickelt und umgesetzt werden.





HUMAN, NOT ALGORITHMIC CONTROL

Die nächste Stufe der Automatisierung entlastet den Menschen, mindert aber seine Entscheidungskompetenz.

Die schnelle und weitreichende Verbreitung Künstlicher Intelligenz hilft, in komplexen oder unübersichtlichen Situationen bessere Entscheidungen zu treffen: von der Entdeckung neuer Musik und Restaurants bis zur Partnerwahl über soziale Netzwerke. Lernende Algorithmen versprechen Einfachheit, Sicherheit und Effizienz. Mit dem wachsenden Potenzial wird das private und berufliche Leben durch diese Möglichkeiten in immer mehr Dimensionen geprägt und ermöglicht, gewonnene Zeit für andere Aufgaben einzusetzen. Die Verlockung der KI-basierten Convenience birgt aber weitreichende Herausforderungen, indem sie die individuelle Entscheidungskompetenz zunehmend unterläuft. Immer mehr Menschen trauen sich nicht mehr zu, eigenständig oder intuitiv zu entscheiden und verlassen sich auf eine algorithmengestützte Empfehlung. Was bei bewusst ausgewählten Lebensbereichen nutzbringend ist, kann sich langfristig negativ auswirken, wenn künftig immer mehr Entscheidungen nicht mehr selbst, sondern durch automatisierte Systeme getroffen – oder vorentschieden – wer-

den. Die Folge ist ein weitreichender Autonomie- und Kontrollverlust, der die Fundamente der Demokratie durch das Einbüßen von Eigenverantwortung unterläuft. Hierin liegt eine wesentlich grössere Gefahr als in inflationär vorgetragenen Szenarien, wo eine künstliche Superintelligenz die Menschheit aufgrund ihrer enormen Denkfähigkeit unterjocht. Mit Blick auf das wachsende Potenzial von smarter Software eröffnen sich hingegen weiterführende Szenarien, in denen intelligente Systeme oder Bots im Auftrag ihrer Auftraggeber Aufgaben, zum Beispiel bei der Identifikation potenzieller Mitarbeitenden, übernehmen. Diese künstlich intelligenten Helfer treffen dabei aber nicht mehr auf die Stellensuchenden selbst, sondern auf deren Bots, die ebenfalls im Auftrag passende Arbeitgeber identifizieren. Im Kontext solcher Bot-Bot-Interaktionen würden die Menschen je nach Perspektive weiter an Einfluss verlieren oder an Freiheitsgraden und Autonomie gewinnen, indem sie sich auf die zentralen zwischenmenschlichen Interaktionen konzentrieren können.





BOOSTING INDIVIDUALS, SHRINKING INDEPENDENCE

Kleinunternehmen und Individuen werden gestärkt, aber gleichzeitig wächst deren Abhängigkeit von intelligenten Systemen.

Die Potenziale Künstlicher Intelligenz führen zu einer Stärkung von Individuen und kleinen Organisationen, die in der Vergangenheit im Umfeld der Economies of Scale an Einfluss oder Marktmacht verloren haben. KI-Systeme bringen dabei eine nächste Dimension der Demokratisierung, die durch den Aufstieg des Internets und die globale Verknüpfung von Menschen durch soziale Netzwerke initiiert wurde. Die Grundlage dafür liefern intelligente Algorithmen oder Robotersysteme, die Berufsleute in ihrer professionellen Entfaltung unterstützen: Ärzte können losgelöst von grossen Spitälern und deren Infrastruktur hochwertige Diagnosen vornehmen, indem KI-Systeme bei der Analyse von Röntgen- oder MRI-Bildern unterstützen. Unabhängige Vermögensverwalter können durch digitale Infrastruktur und Empfehlungssysteme Beratungsleistungen anbieten, die durch die Unterstützung der KI-basierten Cockpitsysteme mit denen von Grossbanken mithalten können. Aber auch Bergbauern und Selbstständige erhalten mehr Autonomie und neue Möglichkeiten durch kostengünstige Unterstützungsleistungen.

Mit Blick auf die nahe Zukunft bedeutet dies, dass die Dichte, aber auch die Qualität von nutzergenerierten Inhalten weiter zunehmen werden – von Produkten bis zu Software, Videos oder anderen kulturellen Erzeugnissen. Dies stellt bestehende Expertisen in unzähligen Berufsgruppen vor neue Herausforderungen, denn ein Musikstück kann mithilfe von KI-gesteuerter Software in kurzer Zeit und ohne teure Infrastruktur erstellt werden. In der Folge gilt es bestehende Definitionen von Qualität in verschiedenen Tätigkeitsbereichen von der Medienproduktion bis zur Beratung oder zum Journalismus zu hinterfragen. Ebenfalls wächst das Risiko einer rückläufigen Systemintegrität und Stabilität durch eine zunehmend vernetzte und algorithmengesteuerte Wertschöpfung. Parallel dazu erhöht die wachsende Leistungsfähigkeit von smarten Maschinen und Algorithmen bei Aufgaben und Prozessen, die mehrheitlich standardisiert ablaufen, den Druck auf Arbeitnehmer, ihre Kompetenzen kontinuierlich anzupassen. Das dürfte den bereits laufenden Strukturwandel im Arbeitsmarkt beschleunigen.





ACCEPTING BLACK BOXES

Neuronale Netzwerke ermöglichen neue Lösungsansätze, sie erfordern aber auch den Umgang mit Intransparenz.

Die mediale Präsenz von künstlich intelligenten Systemen ist in den letzten Jahren von Berichten dominiert worden, bei denen smarte Maschinen Menschen übertreffen – sei es bei Schachspielen, Wissenswettbewerben oder dem Meistern von Knobelaufgaben wie dem Auflösen von Rubik-Würfeln. Bei lernenden Systemen kommen vermehrt Künstliche Neuronale Netzwerke zum Einsatz, die einzelne Mechanismen des menschlichen Gehirns imitieren. Die Analyse der Lösungsvorschläge hat dabei in verschiedenen Fällen gezeigt, dass die künstlich intelligenten Systeme Herausforderungen anders angehen, als es Menschen tun würden und dabei auch neue Lösungskonzepte identifiziert haben, deren Entstehung selbst für die Programmierer der Software nicht mehr nachvollziehbar sind. Mit Blick auf die nahe und fernere Zukunft wird es folglich vermehrt zu KI-basierten Entscheidungen kommen, die für Menschen kaum oder gar nicht nachvollziehbar dargelegt werden

können – aber auch zur Entdeckung von neuen Zusammenhängen und bisher unbekannten Lösungsansätzen in Wissenschaft und Wirtschaft. Hieraus ergeben sich neue Herausforderungen mit Blick auf die wachsende Erwartung an die Transparenz, mit der Algorithmen Entscheidungen oder Kaufempfehlungen abgeben. Insbesondere Regulatoren stehen etwa im Kontext von Finanzdienstleistungen künftig vermehrt vor der Frage, Mechanismen zu bewerten, deren Funktionsweisen nicht nachvollziehbar oder verständlich sind. Daraus entstehen grundlegende Fragen im Umgang mit komplexen Systemen als Grundlage für Innovation oder Entscheidungen in demokratischen Prozessen. Der Zielkonflikt zwischen automatisierter Personalisierung bzw. autonomen Systemen und dem Wunsch nach einem transparenten Einblick in Detailspekte des privaten und gesellschaftlichen Lebens wird Gegenstand einer kontinuierlichen gesellschaftlichen und individuellen Entscheidungsfindung sein.





TOWARDS AN ECONOMY OF TRUST

Simulationen der Realität ermöglichen neue Erlebnisse, liefern aber auch die Basis für Misstrauen.

Die Wahrnehmung der Welt wird zunehmend durch den Einsatz Künstlicher Intelligenz geprägt, die durch Simulationen neue immersive Realitäten erschafft. Diese ermöglichen weitgehend neue Formen von Erlebnissen im Kontext von Computerspielen und der Vermarktung von Produkten, sie führen aber auch zu einem vermehrten sozialen Austausch mit Avataren in Form von Chatbots. Mit den wachsenden Möglichkeiten, die Realität zu simulieren, nehmen aber auch Risiken von Täuschungen und Manipulationen zu. Durch sogenannte Deep Fakes können Bilder oder Filmaufnahmen von Menschen mit wenig Aufwand manipuliert bzw. generiert werden. Die Folge sind Video-Auftritte von Politikern oder Prominenten, deren Lippenbewegungen und Mimik täuschend echt verändert wurden. Oder es entstehen virtuelle Menschen, die nie physisch existieren, aber als Schauspieler oder Influencer eingesetzt werden. Als unvermeidbare Folge wird das Vertrauen in virtuelle Kanäle zurückgehen. Die Suche nach

Authentizitätsgarantien und anderen Formen, die dabei helfen, echte von virtuellen Menschen oder sonstigen Simulationen zu unterscheiden, wird andauern. Als direkte Konsequenz und Gegentrend ist davon auszugehen, dass physische Begegnungen und Interaktionen dort wieder an Bedeutung gewinnen werden, wo es um sensible und private Inhalte geht. Der Aufbau von Vertrauen wird in der Folge zur zentralen Grundlage und Währung für Wirtschaft und Gesellschaft der nahen Zukunft.





EQUALITY IS DEAD. LONG LIVE EQUALITY.

Intelligente Algorithmen können
Fairness steigern, aber ebenso
Diskriminierung verschärfen.

Der Einsatz von lernenden Algorithmen ermöglicht eine effiziente und automatisierte Entscheidungsfindung oder Interaktionen mit intelligenten Schnittstellen. Damit können Informationen nicht nur schneller und günstiger verbreitet werden, sondern es können auch bestehende Vorurteile zwischen Menschen ausgeklammert werden, wenn Algorithmen rein rational entscheiden. Allerdings hat die Realität gezeigt, dass Künstliche Intelligenz zu Diskriminierung und Ausgrenzung führen kann. Bei Auswahlssystemen von Jobkandidaten im HR-Umfeld wurden beispielsweise Menschen mit dunkler Hautfarbe weniger gut bewertet als gleich gut qualifizierte Menschen mit kaukasischen Gesichtszügen. Die Ursache dafür liegt einerseits in Verzerrungen, die in bestehenden Datensätzen abgebildet sind und auf denen die Lernprozesse der Algorithmen aufbauen. Verkürzt gesagt: Algorithmen sind so subjektiv in ihrer Entscheidungsfindung wie die Datengrundlage, mit der sie trainiert werden. Andererseits wurde nachgewiesen, dass

Künstliche Intelligenz bei der Strukturierung von Datensätzen und der Mustererkennung dazu tendiert, Einstellungen und Positionen von Menschen extremer darzustellen, als dass sie es in Realität sind, um homogenere Gruppen und Muster zu erhalten. Mit Blick auf gerechte und faire Entscheidungsgrundlagen, die künftig vermehrt auf KI-basierten Systemen aufbauen werden, sind eine neutrale Datengrundlage und eine Überwachung entlang ethischer Richtlinien durch Menschen eine zentrale Voraussetzung. Ebenso bleibt zu verhandeln, wie lange individuelle Personendaten gespeichert werden dürfen, um darauf basierend Empfehlungen und Prognosen abzugeben. Ein Recht auf Vergessen müsste technologisch definiert werden, um der kulturellen Praxis zu entsprechen.





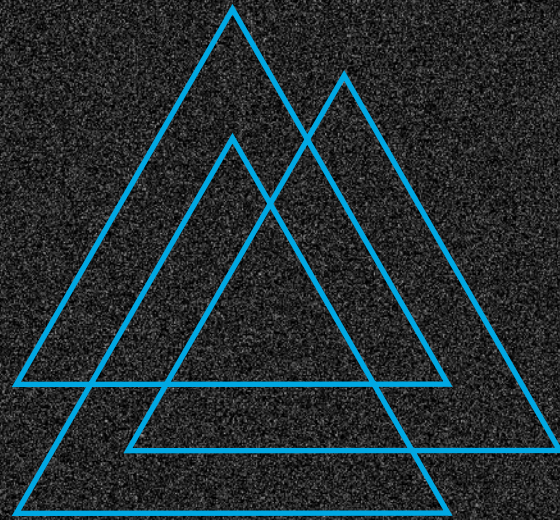
FROM BIG DATA TO SMALL DATA

Die nächste Generation Künstlicher Intelligenz orientiert sich am Gehirn und liefert nicht mehr standardisierte, sondern vielfältige Lösungen.

Der bisherige Schwerpunkt von Künstlicher Intelligenz liegt darin, grosse Datenmengen – Big Data – durch lernende Algorithmen auszuwerten und darauf basierend neue Zusammenhänge zu erkennen oder präzisere Prognosen zu entwickeln. Ein zentrales Forschungsfeld der Zukunft besteht darin, automatisiert Erkenntnisse auch aus kleineren Datenmengen zu gewinnen. Damit sollen KI-Systeme in die Lage versetzt werden – ähnlich wie das menschliche Gehirn –, aus einer überschaubaren Datenmenge Entscheidungsgrundlagen und Rückschlüsse abzuleiten. Ziel wäre, dass ein Algorithmus nicht mehr Hunderttausende Bilder von Katzen analysieren muss, um Katzen auf neuen Bildern zu erkennen, sondern dass nur noch eine geringe Zahl von Inputs für die Wissensgenerierung notwendig ist. Dieser Übergang von Big zu Small Data würde einen Wendepunkt markieren, wie intelligente Maschinen Mehrwerte bringen könnten, da sich damit ihr Einsatz deutlich vereinfachen und flexibilisieren liesse. Bis diese neue Generation jedoch greifbar sein wird, bleiben die absehbaren Vorteile bestehender KI-Systeme gegenüber von starren Algorithmen zentral: Künstliche Intelligenz

ist in der Lage, situativ und flexibel auf neue Herausforderungen zu reagieren und angepasste Lösungen zu entwickeln. Mit Blick auf die immer höhere Komplexität, die sich im Alltag und in allen Sektoren stellt, wären lernende Systeme die Antwort, um mit der zunehmenden Dynamik umzugehen. Im Vergleich zur Vergangenheit, in der Algorithmen homogene und standardisierte Outputs generiert haben, liefern KI-Systeme vielfältige Antworten. Dies erfordert ein Umdenken bei der Definition und Kontrolle von Qualität und den Anforderungen von regulatorischen Prozessen. Unter Berücksichtigung der wachsenden Informationsdichte, die eine gesellschaftsübergreifende Herausforderung darstellt, bekommt Small Data eine doppelte Rolle: Einerseits sollten KI-Algorithmen zum Einsatz kommen, um die Orientierung der Menschen effektiv zu stärken, indem wenige, aber hochwertige Datenpunkte die menschliche Entscheidung bestimmen. Und andererseits sollte das langfristige Ziel für intelligente Systeme darin bestehen, die Welt anhand weniger Datenpunkte interpretieren zu können und damit dem Verständnis von Intelligenz ein echtes Stück näherzukommen.





6

Handlungsfelder

AUFGABEN ZUR GESTALTUNG EINER WÜNSCHBAREN ZUKUNFT

Lange war es Wissenschaftlern und Literaten vorbehalten, die Folgen des Umgangs mit intelligenten Maschinen zu diskutieren. Heute leben wir im Wilden Westen der Künstlichen Intelligenz, mit immer mehr Produkten und Anwendungen, von denen oftmals nicht klar ist, wie sie funktionieren und worin der eigentliche Nutzen besteht. Vor diesem Hintergrund gilt es, Wirtschaft, Politik, Wissenschaft und die Gesellschaft auf die veränderten Rahmenbedingungen vorzubereiten. Ein wesentlicher Teil der Gestaltung der Zukunft betrifft nicht allein die Entwicklung und den Aufbau der technischen Grundlagen, sondern vor allem die Mechanismen einer datenbasierten Wirtschaft und Gesellschaft. Der Fokus von Innovation ist entsprechend holistisch und erfordert bereichsübergreifende Denk- und Umsetzungsstrategien, die regulatorische Basisarbeit ebenso umfassen wie die Weiterentwicklung des Bildungssystems und die Stärkung der gesellschaftlichen Solidarität. Heute eilen die technologischen Fortschritte und die Entwicklung neuer Anwendungen der Gesetzgebung immer mehr voraus. Dies bedeutet nicht, dass für jede neue Herausforderung postwendend eine gesetzliche Reaktion erforderlich ist. Um aber für die Entwicklung und den Einsatz von Künstlicher Intelligenz angemessene Rahmenbedingungen zu schaffen, gilt es frühzeitig und in übergreifenden Netzwerken aus Politik, Wirtschaft und Wissenschaft die Chancen von Künstlicher Intelligenz zu fördern und Risiken zu minimieren. Dies beruht auf der Kopplung der Künstlichen Intelligenz mit der menschlichen.

TECHNISCHE INFRASTRUKTUR AUFBAUEN

Um das Potenzial der Künstlichen Intelligenz zu erschliessen, braucht es Zugang zu hohen Rechenleistungen und Speicherkapazitäten. Diese gilt es mit Blick auf die vielfältigen Anwendungsfelder aufzubauen und einer Vielzahl von Akteuren und Akteurinnen zur Verfügung zu stellen. Damit möglichst vielseitige Mehrwerte entstehen, sollte die grundlegende KI-Infrastruktur nicht nur grossen Technologiekonzernen, sondern auch Organisationen wie Forschungsinstituten, Start-ups und Nichtregierungsorganisationen offenstehen. Dabei können staatliche Institutionen mit Forschungseinrichtungen und privaten Organisationen kollaborieren, um die nötigen finanziellen und technologischen Ressourcen für die Errichtung einer öffentlichen KI-Infrastruktur aufzubringen. Darauf basierend bietet sich die Möglichkeit, in die Inkubationsphase von innovativen Start-ups und Spin-offs zu investieren. Dadurch könnten sich diese vermehrt auch gesellschaftlicher Herausforderungen wie Klimawärmung, Bevölkerungswachstum oder Instabilität des Finanzsystems durch die Entwicklung von KI-basierten Anwendungen annehmen. Durch eine Demokratisierung des Zugangs haben insbesondere auch Individuen die Möglichkeit, auf leistungsfähige Tools zuzugreifen und dadurch Unterstützung bei Entscheidungen zu erhalten oder die eigenen Kompetenzen und die Produktivität zu steigern. Gleichzeitig wird der Wissensaustausch zwischen den

verschiedenen Exponenten gestärkt und dadurch die Erarbeitung innovativer Problemlösungsansätze gefördert.

Darüber hinaus ist eine schnelle und zuverlässige Basisinfrastruktur von zentraler Bedeutung, um KI-Systeme in Echtzeit miteinander kommunizieren zu lassen. In der Mobilität oder in der medizinischen Forschung gehört der Datenaustausch mit minimaler Zeitverzögerung beispielsweise zu den Grundvoraussetzungen, um die Sicherheit der Verkehrsteilnehmer bzw. der Patienten zu gewährleisten. Zu dieser Basisinfrastruktur gehören auch die Aufrechterhaltung und Weiterentwicklung der digitalen Sicherheit im Bereich der Daten, aber vor allem auch bei der Integrität der zunehmend vernetzten Systeme.

DATENBANKEN UND -STRATEGIEN ENTWICKELN

Der Zugang zu qualitativ hochstehenden und zeitlich aktuellen Daten ist die Grundlage für den Einsatz von KI-Systemen und entsprechend von zentraler Bedeutung, um Spitzenforschung im Bereich der Künstlichen Intelligenz zu ermöglichen und nachhaltige Praxisanwendungen zu entwickeln. Heute ist der Zugriff auf strukturierte und hochwertige Datenbanken aufgrund mangelnder Datenqualität sowie durch rechtliche und wirtschaftliche Bestimmungen stark eingeschränkt und fragmentiert.

Eine weitere Grundlage ist der Austausch von Daten mit anderen Akteuren und

«Eine Plattform für informationelle Selbstbestimmung würde uns den Zugriff auf unsere eigenen Daten und damit die Kontrolle über deren Verwendung ermöglichen.»

Dirk Helbing, Professor für Computational
Social Science an der ETH Zürich

Akteurinnen. Dieser setzt aber eine stabile und vertrauenswürdige Infrastruktur voraus. Nur durch eine enge Kollaboration und Koordination zwischen Wirtschaft, Politik und Wissenschaft können die vielseitigen und teilweise divergierenden Interessen der zahlreichen Akteure und Akteurinnen aufeinander abgestimmt und gleichzeitig Persönlichkeitsrechte und die informationelle Selbstbestimmung gewahrt werden. Der Standardisierung von Datenformaten kommt dabei eine zentrale Bedeutung zu. Aufgrund der dezentralen Erfassung und Speicherung von Datensätzen entstehen erhebliche Unterschiede in Bezug auf deren Format, Struktur und Beschriftungsvorgaben. Dies führt oftmals zur Notwendigkeit von aufwendigen und teilweise manuell durchgeführten Bereinigungsverfahren. Eine Harmonisierung der verwen-

deten Standards könnte helfen, hochwertige Daten für Akteure und Akteurinnen aus Wirtschaft, Politik und Wissenschaft über institutionelle und organisatorische Grenzen hinweg nutzbar zu machen.

EIN NEUES NARRATIV FÜR KI ENTWICKELN

Durch einen differenzierten und kritischen Diskurs kann der Begriff der Künstlichen Intelligenz neu verhandelt werden. Das aktuelle Verständnis fokussiert stark auf die Imitation von menschlichen Fähigkeiten als übergeordnetes Ziel der Künstlichen Intelligenz. Dabei kristallisiert sich immer mehr heraus, dass die meisten KI-Systeme der Gegenwart eher als Entscheidungsgrundlage und Unterstützung von professionellen Fachkräften und privaten Nutzern zum Einsatz kommen. Eine realistischere und sachgerechtere Definition des KI-Begriffs ist deshalb eine zentrale Voraussetzung, um sicherzustellen, dass die gemeinsame Vorstellung über Künstliche Intelligenz den realistischen Einsatzgebieten entsprechend – und nicht wie bislang durch die Imagination aus Science-Fiction-Filmen – genährt werden. Dies gründet auch auf der Tatsache, dass die Art und Weise, wie Maschinen ihr Umfeld analysieren, nicht der menschlichen Kognition entspricht. Durch die Neupositionierung von Künstlicher Intelligenz als komplementäre Systeme, die ergänzend zu menschlichen Kompetenzen funktionieren, wird auch den Ängsten vorgebeugt, durch Künstliche Intelligenz ersetzt werden zu können.



BEVÖLKERUNG SENSIBILISIEREN UND VERTRAUEN AUFBAUEN

Mit dem Einzug der Künstlichen Intelligenz in den Alltag der Menschen gilt es, auch in breiten Bevölkerungsschichten ein Verständnis für grundlegende Prinzipien und übergreifende Konzepte der Künstlichen Intelligenz zu erlangen. Dabei geht es weniger um mathematisches oder computerwissenschaftliches Know-how, sondern um die Auseinandersetzung mit Themen wie Mustererkennung, Lernfähigkeit oder den Grenzen der Erklärbarkeit von KI-basierten Ergebnissen. Komplementär dazu muss auch ein Verständnis dafür aufgebaut werden, was Künstliche Intelligenz noch nicht kann: Intuitive Problemlösungen und assoziative Lernverfahren sind nur zwei der Fähigkeiten, durch die sich menschliche Tätigkeiten von den Aufgaben intelligenter Maschinen mittel- bis langfristig abheben werden. Des Weiteren steht die inflationäre Verwendung des KI-Labels im Widerspruch zum Bestreben, Transparenz und Vertrauen zu schaffen. Ein zentraler Aspekt liegt darin, die Bevölkerung im Umgang mit KI-Systemen zu sensibilisieren und sie darüber zu informieren, wann sie mit KI-Systemen interagiert bzw. wie die dahinterstehenden Entscheidungsprozesse funktionieren.

Ein wesentlicher Teil einer solchen Sensibilisierung erfordert auch direkte Berührungspunkte und vor allem Interaktionen von breiteren Bevölkerungsschichten mit künstlich intelligenten Systemen. In Realität finden diese bereits statt, allerdings oft-

mals ohne Bewusstsein der Nutzer. Die aktive Auseinandersetzung mit lernenden Maschinen trägt dazu bei, besser zu verstehen, was Künstliche Intelligenz tatsächlich leistet – und auch wo Grenzen bestehen, beispielsweise beim Versuch, komplexe Gesprächsinhalte mit einem Chatbot zu erörtern. In den USA wird die Bevölkerung aktuell für autonome Fahrzeuge sensibilisiert, indem menschliche Lenker als Autositze verkleidet durch die Strassen fahren und anderen Verkehrsteilnehmern so das Bild eines Fahrzeugs vermitteln, das sicher und ohne Mensch am Steuer navigiert. Eine direkte Interaktion zwischen Künstlicher Intelligenz und dem Alltag der Bevölkerung würde umgekehrt auch dazu beitragen, realistische Lernumgebungen zu gestalten. In solchen Living Labs würden Forschung und Produktentwicklung schneller zu Lösungen gelangen, die für Individuen und Gesellschaft echte Mehrwerte liefern.

BILDUNG UND KOMPE- TENZEN AUF DAS ZEITALTER DER KÜNST- LICHEN INTELLIGENZ AUSRICHTEN

Der Einsatz von Künstlicher Intelligenz wird die Anforderungen im Arbeitsumfeld auf mehreren Ebenen verändern. Zunächst entstehen neue Berufsbilder, die wiederum neue Kompetenzen erfordern. Beispielsweise ist vorstellbar, dass Trainer von KI-Systemen die algorithmischen Lernprozesse überwachen und analysieren, Interpretieren die Ergebnisse von automatisierten

«Es wäre sinnvoll, wenn man bei der technischen Ausbildung Problemstellungen im Bereich der KI auch vermehrt aus einer ethischen und politischen Perspektive betrachtet.»

Gabriel Krummenacher, PhD in Machine Learning und
Lead Data Scientist bei Zühlke Engineering AG

Entscheidungssystemen erklären oder Ethiker die Richtlinien für KI-basierte Prozesse definieren. Es bleibt jedoch festzuhalten, dass niemand heute weiss, wie die Berufe der Zukunft im Detail aussehen werden. Klar ist jedoch, dass die erforderlichen Kompetenzen massgeblich durch Agilität und Anpassungsfähigkeit bestimmt werden, was insbesondere die Relevanz einer begleitenden Weiterbildung in sämtlichen Lebensabschnitten erhöht. Darüber hinaus werden künstlich intelligente Systeme zu einer fortschreitenden Personalisierung und Virtualisierung der Wissensvermittlung beitragen. Dabei gilt es neue Impulse aus dem sogenannten Education-Tech-Sektor aufzunehmen und dadurch auf allen Bildungsebenen von Grund- bis zu Hochschulen und unternehmensspezifischen Weiterbildungen virtuelle, situative und personalisierte Wissensvermittlung voranzutreiben. Dies wird auch die räumliche Organisation von Bildungseinrichtungen massgeblich beeinflussen: Klassische Hörsäle für Hunderte von Zuhörern dürften zugunsten von digitalen Vermittlungsfor-

maten an Bedeutung verlieren. Gleichzeitig wird der persönliche Austausch in kleinen Gruppen an Relevanz gewinnen. Im Fokus neuer Lösungen werden entsprechend Schnittstellen zwischen digitalen und persönlichen Vermittlungsformaten stehen. Letztlich sollten ethische Fragestellungen im Hinblick auf die gesellschaftlichen Auswirkungen von KI-Systemen vermehrt in die technische Ausbildung von Programmierern und Systementwicklern miteinfließen.

POSITIONIERUNG IM INTERNATIONALEN WETTBEWERB STÄRKEN

Künstliche Intelligenz ist längst zu einem zentralen Wirtschaftstreiber geworden. Entsprechend ist die Qualität von Forschung und Entwicklung auf dem Gebiet der Künstlichen Intelligenz ein Differenzierungsfaktor in der globalen, digitalen Welt geworden. China strebt mit einer enormen Dynamik und hohen Investitionen um die Vormachtstellung als erste internationale

KI-Supermacht – neben den USA. Innerhalb dieses Wettlaufs geht es nicht allein um marktwirtschaftliche Dominanz. Gerade hinsichtlich der unterschiedlichen Wertesysteme einer noch immer liberalen Marktwirtschaft in den USA und dem autokratischen System Chinas werden sich die Potenziale der Künstlichen Intelligenz zunehmend auch auf die gesellschaftlichen Strukturen auswirken. Um im internationalen Wettbewerb relevant zu bleiben, kann der Schlüssel darin liegen, sich auf ein bestimmtes Teilgebiet der Künstlichen Intelligenz zu fokussieren und dadurch Doppelspurigkeiten zu vermeiden. Kein Standort kann an allen Fronten der KI-Forschung eine weltweite Spitzenposition einnehmen. Ein klarer Fokus erlaubt es, die benötigten Kräfte aus Politik, Wirtschaft und Wissenschaft zu bündeln. Verschiedene Faktoren wie wirtschaftliche Voraussetzungen, systemische Unterschiede oder die spezifische Akzeptanz in der Bevölkerung bestimmen dabei, welches Teilgebiet für den jeweiligen Standort infrage kommt. Sobald der Fokus auf ein spezifisches Teilgebiet gelegt wurde, können gezielt Fachkräfte aus den relevanten Bereichen rekrutiert und langfristig an den Standort gebunden werden. Experten im Bereich der Künstlichen Intelligenz sind umkämpft. Deshalb gilt es insbesondere die Rahmenbedingungen und Migrationsbestimmungen für Experten kritisch zu überprüfen, um die eingeschlagene KI-Strategie mit der nötigen Konsequenz voranzutreiben. Die Rolle von Europa und der Schweiz kann in dieser Suche nach einer Positionierung darin liegen, einen dritten Weg für die Nutzung von Künstlicher Intelligenz aufzu-

zeigen, der neben der technischen Exzellenz den gesellschaftlichen Nutzen ins Zentrum rückt, ohne die Grund- und Menschenrechte infrage zu stellen. Gerade mit Blick auf den verstärkten KI-Forschungsfokus, abseits des Big-Data-Ansatzes aus kleinen Datenmengen intelligente Entscheidungen abzuleiten, ergibt sich eine klare Chance, um auch ohne Zugang zu grossen Absatzmärkten mit Hunderten Millionen Menschen wie den USA oder China weiterhin international eine führende Rolle einzunehmen.

GESELLSCHAFTLICHE UND ETHISCHE GRUNDLAGEN KLÄREN

Der Einsatz von intelligenten Algorithmen zur Aufbereitung von Informationen oder zur Segmentierung von Kunden und Kundinnen legt die Basis für neue Geschäftsmodelle und Organisationsformen. Beispielsweise ermöglichen dynamische Preise, flexibel auf variierende Marktgegebenheiten zu reagieren, sie bringen aber auch Herausforderungen mit sich, wie die Diskriminierung einzelner Bevölkerungsgruppen als Folge von verzerrten Datensätzen, die als Lerngrundlage eingesetzt werden. Um diese Risiken zu limitieren oder einzugrenzen, wächst die Notwendigkeit für ethische Grundlagen. Aktuell entwickeln öffentliche und Non-Profit-Organisationen, aber auch kommerzielle Unternehmen Chartas als Basis für den Umgang mit datenbasierten oder KI-gestützten Geschäftsmodellen. Diese betreffen Fragen des Datenbesitzes, der je nachdem, ob die Daten individuell oder kollektiv ausgewer-

tet werden, unterschiedlich ausfallen kann. Des Weiteren werden Fragen, wie und ob die Entscheidungsmechanismen von Algorithmen transparent zugänglich gemacht werden sollen, geklärt. Im grösseren Bild betreffen diese gesellschaftlichen Grundlagen nicht allein Experten und Organisationen, es braucht auch einen gesellschaftlichen Konsens darüber, wo und wie künstlich intelligente Systeme tatsächlich wünschenswert sind. Im Zentrum der ethischen Grundlagen stehen Transparenz, Gleichberechtigung, Sicherheit und Selbstbestimmung. Dazu müssen die Akteure und Akteurinnen aus Wirtschaft, Politik und Wissenschaft gemeinsame Richtlinien und Standards erarbeiten, die auch die lokalen kulturellen Umstände berücksichtigen. Der Diskurs sollte auch sensible Themen wie die Etablierung einer neuen Fehlerkultur im Umgang mit KI-basierten Fehlleistungen aufnehmen. Durch die Entscheidungsfähigkeit von KI-Systemen entsteht die Gefahr, dass Menschen psychische oder physische Schäden davontragen, ohne dass eine Person dafür verantwortlich gemacht werden kann. Dies ist bei automatisierten Rekrutierungsentscheidungen

ebenso relevant wie im Kontext der teilautonomen Steuerung von Fahrzeugen im Strassenverkehr. Juristische Fragen nach der Haftung und Verantwortlichkeit bei KI-basierten Fehlern sollten deshalb frühzeitig in den gesellschaftlichen Diskurs aufgenommen werden.

Mit Blick auf eine allfällige weitere signifikante Zunahme der Leistungsfähigkeit Künstlicher Intelligenz fordern Experten für eine fernere Zukunft umgekehrt auch Grundrechte für Künstliche Intelligenz, wenn sich diese den Verhaltensweisen eines Menschen weiter annähern sollte. Aus heutiger Sicht ist dies noch in weiter Ferne; als Teil der grundsätzlichen Auseinandersetzung bei der Suche danach, was Menschen auszeichnet, sind solche Gedankenspiele jedoch hilfreich.

ÖKOSYSTEME AUFBAUEN

Mit Blick auf die vielfältigen Chancen und komplexen Herausforderungen der daten- und KI-basierten Welt ist es naheliegend, dass eine übergreifende Zusammenarbeit

«Die Verzahnung zwischen Industrie und Akademie ist extrem wichtig, um einen KI-Standort international konkurrenzfähig zu machen.»

Thomas Hofmann, Professor für Machine Learning an der ETH Zürich

zwingend ist, um Potenziale zu stärken und Risiken zu limitieren. Die Schaffung von transdisziplinären und breit abgestützten Netzwerken ist entscheidend, um die Position von Wirtschafts- und Forschungsstandorten der Künstlichen Intelligenz zu stärken. Die geografische Nähe zwischen den Akteuren und Akteurinnen aus Politik, Wirtschaft und Wissenschaft kann dabei ein Vorteil sein, um Cluster mit KI-Fokus zu bilden und von der vielseitigen Expertise zu profitieren. Eine branchenübergreifende Zusammenarbeit wird für die Mehrheit aller Sektoren und Clusters zur Grundlage, nicht nur wenn es um die Entwicklung von digitalen oder KI-basierten Lösungen geht, sondern gerade um dem viel geforderten Wunsch einer echten Kundenzentrierung gerecht zu werden. Das bessere Verständnis der individuellen Bedürfnisse trägt dazu bei, diese neuen Netzwerke besser aufzubauen und sie zu orchestrieren. Gleichzeitig steigt die Komplexität dieser Vorhaben, da mit dem Fokus auf individuelle Bedürfnisse auch keine vorgegebenen, sondern spezifische und vor allem flexible Kooperationen zur Voraussetzung werden. Entsprechende KI-Ökosysteme stellen Organisationen vor eine Vielzahl von Herausforderungen – von der Identifizierung der relevanten künftigen Bedürfnisse, die als Grundlage für die Partnerschaften dienen, bis zur Gestaltung der daten- und tendenziell echtzeitbasierten Strukturen. Innerhalb dieser dynamischen Netzwerke gilt es sich zu positionieren und Geschäftsmodelle zu definieren. Vor allem aber erfordern agile Netzwerke mehr denn je eine langfristige Perspektive als Fluchtpunkt – und

einen Orchestrator mit Verantwortung für einheitliche Zielvorgaben.

NEUE AUFGABEN FÜR DIE ÖFFENTLICHE HAND DEFINIEREN

Die durch den Einsatz von Künstlicher Intelligenz veränderten Rahmenbedingungen bestimmen auch die Aufgaben von öffentlichen Institutionen. Eine systematische und frühzeitige Auseinandersetzung mit den möglichen Ausprägungen, Chancen und Herausforderungen wird zur Grundvoraussetzung für Politik und Verwaltung von Gemeinden bis auf Bundesebene. Dies legt die Basis für eine smarte Regulierung, die wichtige Themen aufgreift, bevor diese den Alltag der Bürgerinnen und Bürger massgeblich beeinflussen. Damit können kritische Themen auch in einen öffentlichen Dialog gebracht werden, was einer punktuellen und einseitigen Reaktion vorgreift. Der Übergang in eine KI-basierte Wirtschaft und Gesellschaft hat dabei aber noch weitere Auswirkungen auf die regulatorischen Prozesse. Grundsätzlich ist davon auszugehen, dass die Vielfalt von Lösungsansätzen von Arbeits- bis zu Geschäftsmodellen signifikant zunehmen wird. Die daraus resultierende Komplexität erschwert die Steuerung durch traditionelle Kontrollmechanismen, weil sich die hohe Dynamik und die vielfältigen Folgen kaum mehr durch einige wenige Richtlinien handhaben lassen werden. Ob schon ein Aufstocken des Leistungsvolumens eine naheliegende Lösung scheint, dürfte sie der Realität nicht gerecht werden. Die Antwort auf die Komplexität und

«Die Regulierung kann zurzeit nicht mit der KI-Entwicklung Schritt halten. Deshalb ist es sinnvoll, kontrolliert zu experimentieren und erst dann entsprechende Gesetze zu erlassen.»

Jana Koehler, Informatikprofessorin und wiss.-techn.
Geschäftsführerin des Deutschen Forschungszentrums für KI

Vielfalt der digitalen oder KI-basierten Welt ist deshalb notwendigerweise, die Spielräume für neue Lösungen zu erweitern und diese dennoch mit klaren Grenzen und Werterichtlinien zu definieren. In diesem Sinn könnten Unternehmen mit Politik und Wissenschaft unter Einbezug der Zivilgesellschaft vermehrt gestaltende rechtliche Gefässe schaffen, in denen unter strenger Beaufsichtigung kontrolliert experimentiert wird, um erste Erfahrungen zu sammeln und darauf basierend entsprechende Gesetze und Verordnungen zu erlassen. Aus regulatorischer Perspektive wird eine weitere Aufgabe darin bestehen, Transparenz hinter den Entscheidungsprozessen von KI-Systemen zu gewährleisten und gleichzeitig das geistige Eigentum von Firmen zu schützen. Falls von Firmen verlangt wird, ihre Algorithmen offenzulegen, kann dies zu einer Hemmung der Innovationskraft führen, da sich der Aufwand für die Forschung und Entwicklung von KI-Systemen langfristig nicht mehr auszahlt. Dies würde der Idee, vielfältige Lösungsansätzen zu

schaffen und für breite Bevölkerungsschichten Mehrwerte zu generieren, entgegenlaufen. Praxistaugliche regulatorische Antworten zu diesem Spannungsfeld können deshalb einen entscheidenden Standortvorteil darstellen.

Aufgrund der globalen Dimensionen des Einsatzes von Künstlicher Intelligenz kann eine wirksame Regulierung theoretisch nur teilweise lokal erfolgen. Übergreifende und supranationale Strukturen gewinnen entsprechend an Bedeutung, unterliegen aber denselben Limitationen im Umgang mit Komplexität, die sich in einem globalen Kontext noch weiter akzentuieren. Folglich wächst vor allem die Bedeutung von nationalen und regionalen Regelwerken und Richtlinien als Vorlage für eine Übertragung in grössere geografische Dimensionen. Den Gemeinden und Kantonen kommt eine zentrale Rolle zu, um als Gestalter in der Bestimmung von Richtlinien und Verhaltensweisen im Umgang mit Künstlicher Intelligenz eine Vorreiterrolle einzunehmen.

W.I.R.E. ist ein führender interdisziplinärer Think Tank, der sich seit rund zehn Jahren mit globalen Entwicklungen in Wirtschaft, Wissenschaft und Gesellschaft beschäftigt. Im Fokus des Schweizer Denklabors stehen die frühe Erkennung neuer Trends und deren Übersetzung in strategische Landkarten und Handlungsfelder für Unternehmen und öffentliche Institutionen.

An der Schnittstelle zwischen Wissenschaft und Praxis zeichnet sich W.I.R.E. durch eine kritische Denkhaltung und politische Neutralität aus. Thematische Schwerpunkte betreffen die digitale Wirtschaft, gesellschaftliche Innovation und die Förderung der Zukunftsfähigkeit. Dabei stellt der Think Tank seine Expertise in den Dienst von Öffentlichkeit, Unternehmen und Behörden – von Life Sciences, Finanzdienstleistern und Medien bis hin zu Food und Industrie.

Die zwei- und dreidimensionale Wissensvermittlung von W.I.R.E. zeichnet sich durch die Verbindung von Inhalt und Form sowie einen hohen Anspruch an Ästhetik und Design aus. Der Think Tank verfügt über ein internationales Netzwerk aus Experten, Vordenkern und Entscheidungsträgern.

www.thewire.ch

Die Standortförderung ist Ansprechpartnerin bei Fragen zur Firmengründung und Standortwahl. Zwecks Stärkung zukunftsweisender Wirtschaftszweige wie Cleantech, Finance, Life Sciences, Kreativwirtschaft, Informations- und Kommunikationstechnologie fördert sie die Vernetzung von Unternehmen und Institutionen entlang der Wertschöpfungskette im Raum Zürich. Die Standortförderung ist aber auch Anlaufstelle für Anliegen von ansässigen Unternehmen oder bei Bewilligungsfragen. Ausserdem engagiert sie sich für die administrative Entlastung der Unternehmen.

Kurz: Die Standortförderung begleitet, beschleunigt und vernetzt als Bindeglied zwischen Wirtschaft und öffentlicher Verwaltung.

www.standort.zh.ch

www.zh.ch/standort [ab Mai 2020]

EXPERTEN

DIRK HELBING

Professor für Computational Social Science an der ETH Zürich

THOMAS HOFMANN

Professor für Machine Learning an der ETH Zürich

PASCAL KAUFMANN

Neurowissenschaftler und KI-Unternehmer

JANA KOEHLER

Informatikprofessorin und wiss.-techn. Geschäftsführerin des Deutschen Forschungszentrums für KI

GABRIEL KRUMMENACHER

PhD in Machine Learning und Lead Data Scientist bei Zühlke Engineering AG

EVA WEBER-GUSKAR

Philosophieprofessorin und Fellow der Digital Society Initiative der Universität Zürich

AUTOREN

RAPHAEL VON THIESSEN

Raphael von Thiessen ist Projektleiter und Researcher bei W.I.R.E. und analysiert im Rahmen von strategischen Projekten den Einfluss von Trends und Technologien auf Wirtschaft, Wissenschaft und Gesellschaft. Nach seinem Bachelor in Politikwissenschaft an der Universität Zürich setzte er digitale Innovationen für das Medienunternehmen Axel Springer um. Anschliessend absolvierte er an der ESCP Europe in London, Paris und Berlin einen Master of Science in Betriebswirtschaft und verfasste seine Masterarbeit zur Entwicklung von Geschäftsmodellen, die auf Technologien der Künstlichen Intelligenz basieren. Zuletzt war er als Unternehmensberater mit dem Themenschwerpunkt Future of Work für verschiedene Firmen tätig.

STEFAN PABST

Stefan Pabst ist Senior Projektleiter beim Think Tank W.I.R.E. Seit seinem Studium der Philosophie und Physik betreut er Innovations- und Strategieprojekte innerhalb der Themenfelder Digitalisierung und gesellschaftliche Innovation. Ergänzt durch seine Expertise in der interdisziplinären Ideenfindungsmethode «Design Thinking» begleitet er Organisationen beim Erschliessen neuer Marktfelder und bei der aktiven Gestaltung eines digitalen Kulturwandels. Er ist mitverantwortlich für eine Vielzahl von Publikationen des Think Tanks und tritt als Keynote-Referent auf. Seine Themenschwerpunkte liegen in der Transformation von Kommunikation, Mobilität und gesellschaftlichem Zusammenhalt sowie in den ethischen Fragestellungen im Kontext der Automatisierung.

DR. STEPHAN SIGRIST

Stephan Sigrist ist Gründer und Leiter des Think Tanks W.I.R.E. Er analysiert seit vielen Jahren interdisziplinär Entwicklungen in Wirtschaft, Wissenschaft und Gesellschaft und beschäftigt sich schwergewichtig mit den Folgen der Digitalisierung in den Life Sciences, Financial Services, Medien, Infrastruktur und Mobilität. Er ist Herausgeber der Buchreihe ABSTRAKT und Autor zahlreicher Publikationen sowie Keynote-Referent an internationalen Tagungen. Mit W.I.R.E. unterstützt er Entscheidungsträger bei der frühen Erkennung relevanter übergeordneter Entwicklungen und deren Übersetzung in langfristige Strategien und Innovationsprojekte.

QUELLENVERZEICHNIS

EINLEITUNG

1 Minds, brains, and programs. Behavioral and Brain Sciences 3 (3): 417–457. John. R. Searle, (1980). <http://cogprints.org/7150/1/10.1.1.83.5248.pdf> (abgerufen am 16.10.2019).

2 Artificial Intelligence: A modern approach (third edition). Stuart Russell and Peter Norvig (2010).

3 Computing Machinery and Intelligence. Mind 49: 433–460. A. M. Turing (1950). <https://www.csee.umbc.edu/courses/471/papers/turing.pdf> (abgerufen am 18.10.2019).

4 Computer Chess History. Bill Wall (1917). https://www.chessprogramming.org/Herbert_Simon (abgerufen am 15.10.2019).

5 Cramming More Components onto Integrated Circuits. Gordon Moore (1965). <http://ai-media-companies-hide-likes-follower-numbers-Twitter-rant.html>. (abgerufen am 23.9.2018).

TECHNOLOGISCHE GRUNDLAGEN

1 The Pattern Recognition Basis of Artificial Intelligence. Donald R. Tvetter (1998). <http://www.dontveter.com/basisofai/basisofai.html> (abgerufen am 12.10.2019).

2 Information Theory: Physiology. Encyclopaedia Britannica (2019). <https://www.britannica.com/science/information-theory/Physiology> (abgerufen am 14.10.2019).

3 How we might protect ourselves from malicious AI. MIT Technology Review. Karen Hao (2019). https://www.technologyreview.com/s/613555/how-we-might-protect-ourselves-from-malicious-ai/?utm_medium=tr_social&utm_campaign=site_visitor.unpaid.engagement&utm_LinkedIn#Echobox=1568732099 (abgerufen am 15.10.2019).

4 Know your Adversary: understanding Adversarial Examples. Cody Marie Wild (2019). <https://towardsdatascience.com/know-your-adversary-understanding-adversarial-examples-part-1-2-63af4c2f5830> (abgerufen am 16.10.2019).

5 Tschacher & Storch (2012). Die Bedeutung von Embodiment für Psychologie und Psychotherapie. Psychotherapie in Psychiatrie, Psychosomatischer Medizin und Klinischer Psychologie 17. Tschacher & Storch (2012).

6 Building Brains for Bodies, Autonomous Robots 1 (1). Rodney Brooks und Lynn Andrea Stein (1994). <https://link.springer.com/article/10.1007/BF00735340> (abgerufen am 14.10.2019).

7 Pattern Recognition and Natural Language Processing: State of the Art. TEM Journal 5 (2). Mirjana Kocaleva et al. (2016). http://www.temjournal.com/content/52/TemJournalMay2016_236_240.pdf (abgerufen am 20.10.2019).

8 Conceptual Biology, Hypothesis Discovery, and Text Mining: Swanson's Legacy. Tania

Bekhuis (2006). https://www.researchgate.net/publication/7194619_Conceptual_biology_hypothesis_discovery_and_text_mining_Swanson's_legacy (abgerufen am 13.10.2019).

9 The Relationship between Learning and Intelligence. Learning and Individual Differences 1 (1). Arthur R. Jensen (1989). <https://www.sciencedirect.com/science/article/pii/1041608089900095> (abgerufen am 9.10.2019).

10 Some studies in Machine Learning using the Game of Checkers. IBM Journal of Research and Development 44 (1.2). A.L. Samuels (2000). <https://ieeexplore.ieee.org/abstract/document/5389202> (abgerufen am 8.10.2019).

11 Machine Learning for Medical Imaging. RadioGraphics 2017 (37) Bradley J. Erickson et al. (2017). <https://pubs.rsna.org/doi/pdf/10.1148/rg.2017160130> (abgerufen am 14.10.2019).

12 Real World Applications of Unsupervised Learning. Ashwin Joy (2019). <https://pythonistaplanet.com/applications-of-unsupervised-learning/> (abgerufen am 13.10.2019).

13 AlphaGo Zero: Starting from Scratch. David Silver und Demis Hassabis (2017). <https://deepmind.com/blog/article/alphago-zero-starting-scratch> (abgerufen am 14.10.2019).

14 The truth behind Facebook inventing a new language. Roman Kucera (2017). <https://towardsdatascience.com/the-truth-behind-facebook-inventing-a-new-language-37c5d680e5a7> (abgerufen am 16.10.2019).

15 Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. Yonghui Wu et al. (2016). <https://arxiv.org/abs/1609.08144> (abgerufen am 20.10.2019).

16 Deep Learning for the Classification of Lung Nodules. He Yang (2016). <https://arxiv.org/ftp/arxiv/papers/1611/1611.06651.pdf> (abgerufen am 12.10.2019).

17 Evolving simple programs for playing Atari games. Wilson et al. (2018). <https://www.technologyreview.com/s/611568/evolutionary-algorithm-outperforms-deep-learning-machines-at-video-games/> (abgerufen am 20.10.2019).

18 Maschinelles Lernen: Korrelationen müssen auch kausal sein. Martin Spindler, Victor Chernozhukov und Ye Luo (2019). <https://www.faz.net/aktuell/wirtschaft/diginomics/maschinelles-lernen-korrelationen-muessen-auch-kausal-sein-16239535.html> (abgerufen am 15.10.2019).

19 What we can learn from the epic Failure of Google Flu Trends. David Lazer und Ryan Kennedy (2015). <https://www.wired.com/2015/10/can-learn-epic-failure-google-flu-trends/> (abgerufen am 14.10.2019).

20 English leads in Speech Recognition, but not for long.

Bryan Healey (2016). <https://medium.com/@lola.com/english-leads-in-speech-recognition-but-not-for-long-b21893bd2fcf> (abgerufen am 10.10.2019).

21 Ongoing Face Recognition Vendor Test. National Institute of standards and technology. U.S. Department of Commerce (2019). https://www.nist.gov/sites/default/files/documents/2019/07/03/frvt_report_2019_07_03.pdf (abgerufen am 11.10.2019).

22 How IBM Watson overpromised and underdelivered on AI Health Care. Eliza Strickland (2019). <https://spectrum.ieee.org/biomedical/diagnostics/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care> (abgerufen am 27.10.2019).

23 Image Classification using Deep Neural Networks – A Beginner friendly Approach using TensorFlow. Abdellatif Abdelfattah (2017). <https://medium.com/p/94b0a090ccd4/responses/show> (abgerufen am 12.10.2019).

24 Deep Learning (Adaptive Computation and Machine Learning Series). Ian Goodfellow, Yoshua Bengio und Aaron Courville (2019).

25 Federated Learning: Collaborative Machine Learning without Centralized Training Data. Brendan McMahan und Daniel Ramage (2017). <https://ai.googleblog.com/2017/04/federated-learning-collaborative.html> (abgerufen am 13.10.2019).

26 Designing and implementing Transparency for real time Inspection of autonomous Robots. Connection Science 29 (3). Andreas Theodorou et al. (2017). <https://www.tandfonline.com/doi/abs/10.1080/09540091.2017.1310182> (abgerufen am 12.10.2019).

27 How AI Detectives are cracking open the black box of deep learning. Paul Voosen (2017). <http://www.sciencemag.org/news/2017/07/how-ai-detectives-are-cracking-open-black-box-deep-learning> (abgerufen am 12.10.2019).

28 Energy and Policy Considerations for Deep Learning in NLP. Emma Strubell, Ananya Ganesh und Andrew McCallum (2019). <https://arxiv.org/abs/1906.02243> (abgerufen am 12.10.2019).

ANWENDUNGSBEREICHE

1 Fast and accurate View Classification of Echocardiograms using Deep Learning. Nature Research Journal 1 (6). Ali Madani et al. (2018). <https://www.nature.com/articles/s41746-017-0013-1> (abgerufen am 12.10.2019).

2 Unilever saves on Recruiters by using AI to assess Job Interviews. Robert Booth (2019). <https://www.theguardian.com/technology/2019/oct/25/unilever-saves-on-recruiters-by-using-ai-to-assess-job-interviews> (abgerufen am 29.10.2019).

3 The Accuracy, Fairness, and Limits of predicting Recidivism. Julia Dressel und Hany Farid (2018). <https://advances.sciencemag.org/content/4/1/eaao5580/tab-pdf> (abgerufen am 22.10.2019).

4 Künstliche Intelligenz verbessert Wettervorhersagen. Nicole Lücke (2019). <https://www.ingenieur.de/technik/fachbereiche/ittk/kuenstliche-intelligenz-verbessert-wettervorhersagen/> (abgerufen am 13.10.2019).

5 Treating Addiction with an App. Nanette Byrnes (2017). <https://www.technologyreview.com/s/604085/treating-addiction-with-an-app/> (abgerufen am 22.10.2019).

6 Your next Doctor's Appointment might be with an AI. Douglas Heaven (2018). <https://www.technologyreview.com/s/612267/your-next-doctors-appointment-might-be-with-an-ai/> (abgerufen am 13.10.2019).

7 Perceptions of Chatbots in Therapy. Samuel Bell, Clara Wood und Advait Sarkar (2019). https://www.microsoft.com/en-us/research/uploads/prod/2019/03/bell_2019_chatbots-camera-ready.pdf (abgerufen am 28.10.2019).

8 Apple veröffentlicht iPhone 4S, iOS 5 & iCloud. Apple (2011). <https://www.apple.com/chde/newsroom/2011/10/04Apple-Launches-iPhone-4S-iOS-5-iCloud/> (abgerufen am 13.10.2019).

9 Meet the Robot that's hiring Humans for some of the World's major Corporations <https://www.cnn.com/2018/04/20/this-robot-hires-humans-for-some-major-corporations.html> (abgerufen am 12.10.2019).

10 Artificial Intelligence: A modern Approach (third edition). Stuart Russell and Peter Norvig (2010).

11 A Deep Dive on Robotic Technology & Virtual Assistants. AI Med Magazine (5). Tamás Haidegger (2019). <https://ai-med.io/artificial-intelligence-robotic-surgery/> (abgerufen am 10.10.2019).

12 Curiosity Rover gets a Boost from Artificial Intelligence (2017). <https://www.nature.com/articles/d41586-017-00626-6> (abgerufen am 11.10.2019).

13 A smart artificial Hand for Amputees merges User and Robotic Control (2019). <https://actu.epfl.ch/news/a-smart-artificial-hand-for-amputees-merges-user-a/> (abgerufen am 8.10.2019).

14 Celebrating Johann Sebastian Bach. <https://www.google.com/doodles/celebrating-johann-bach> (abgerufen am 12.10.2019).

15 Understanding Artificial Intelligence: Fundamentals, Use Cases and Methods for a Corporate AI Journey. Marie Sirrenberg and Ralf Kreutzer (2019).

16 This is Speed Gate: The world's first Sport imagined by an AI. <https://playspeedgate.org> (abgerufen am 13.10.2019).

17 What is a Deep Fake: The Economist explains. Susan Sontag (2019). <https://www.economist.com/the-economist-explains/2019/08/07/what-is-a-deepfake> (abgerufen am 13.10.2019).

KONSEQUENZEN

1 Collaborative Intelligence: Humans and AI are joining Forces. James Wilson and Paul Daugherty (2018). <https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces> (abgerufen am 8.10.2019).

2 How Retailers can keep up with Consumers. Ian McKenzie, Chris Meyer and Steve Noble (2013). <https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers> (abgerufen am 10.10.2019).

3 Ten Ways autonomous driving could redefine the automotive World. Michele Bertonecello and Dominik Wee (2015). <https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/ten-ways-autonomous-driving-could-redefine-the-automotive-world> (abgerufen am 13.10.2019).

4 AlphaGo Zero: Starting from Scratch. David Silver and Demis Hassabis (2017). <https://deepmind.com/blog/article/alphago-zero-starting-scratch> (abgerufen am 14.10.2019).

5 AI Analysis: Sizing the Prize. Dr. Anand Rao und Gerard Verweij (2019). <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf> (abgerufen am 3.11.2019).

6 High Level Expert Group on Artificial Intelligence. European Commission (2019). <https://ec.europa.eu/digital-single-market/>

[en/high-level-expert-group-artificial-intelligence](https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence) (abgerufen am 12.10.2019).

7 Amazon scraps secret AI recruiting tool that showed bias against women. Jeffrey Dastin (2018). <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G> (abgerufen am 13.10.2019).

8 The uncanny Valley. IEEE Robotics & Automation Magazine 19 (2). Masahiro Mori (2012). <https://ieeexplore.ieee.org/document/6213238> (abgerufen am 13.10.2019)

9 Artificial Intelligence: A modern Approach (third edition). Stuart Russell and Peter Norvig (2010).

10 Complications: A Surgeon's Notes on an Imperfect Science. Atul Gawande (2002).

11 The next Privacy Hurdle: Teaching AI how to forget. Darren Shou (2019). <https://www.wired.com/story/the-next-big-privacy-hurdle-teaching-ai-to-forget/> (abgerufen am 15.10.2019).

12 Our ongoing work to tackle hate. Youtube (2019). <https://youtube.googleblog.com/2019/06/our-ongoing-work-to-tackle-hate.html> (abgerufen am 28.10.2019).

IMPRESSUM

© 2020 W.I.R.E. und NZZ Libro, Schwabe Verlagsgruppe AG, Basel
www.thewire.ch

Im Auftrag der Standortförderung im Amt für Wirtschaft und Arbeit des Kantons Zürich

Konzept: W.I.R.E.

Inhalte: Raphael von Thiessen, Stefan Pabst, Leonie Mathis, Dr. Stephan Sigris

Gestaltung, Illustration: Sibylle Brodbeck, Swati Prasad

Korrektur: Ruth Rybi

Druck: Neidhart + Schön Print AG, Zürich

Herausgabe: Februar 2020

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werks oder von Teilen dieses Werks ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechts.

Diese Publikation von W.I.R.E. erscheint in Deutsch.

ISBN 978-3-03810-439-1

www.nzz-libro.ch

NZZ Libro ist ein Imprint der Schwabe Verlagsgruppe AG

Das Zeitalter der Künstlichen Intelligenz kündigt sich an. Was lange Zeit von Science-Fiction-Autoren vorgezeichnet wurde, scheint heute Realität zu werden: Intelligente und lernende Maschinen, die im Umfeld von Banken bis zur Medizin selbstständig Aufgabenstellungen lösen und den Menschen anstrengende Aufgaben abnehmen. Doch der Ausblick ist nicht nur optimistisch. Skeptische Stimmen befürchten die baldige Machtübernahme einer Superintelligenz, Massenarbeitslosigkeit oder die Dystopie eines digitalen Überwachungsstaats.

Schnelle Fortschritte und grosse Ankündigungen erhöhen aktuell die Komplexität: Die Übersicht zu behalten und das reale Potenzial von künstlichen intelligenten Systemen einzuschätzen wird schwieriger. In dieser Situation sind kritisches Denken und eine ganzheitliche Perspektive auf den technologischen Wandel wichtiger denn je. Doch welche Chancen bietet Künstliche Intelligenz tatsächlich? Welche Versprechungen der Künstlichen Intelligenz sind Hypes, auf welche gilt es sich vorzubereiten? Was sind die Risiken, die autonome Entscheidungs- oder Steuerungssysteme mit sich bringen? Und welche strategischen Handlungsfelder ergeben sich daraus für Unternehmen, die Wissenschaft, aber auch für Politik und Verwaltungen?

Um diese Fragen zu beantworten, braucht es ein differenziertes Verständnis der technischen Grundlagen und das Ableiten von mittel- bis langfristigen Folgen für Unternehmen, Bürger und die Gesellschaft. Der Think Tank W.I.R.E. hat sich in Zusammenarbeit mit dem Amt für Wirtschaft und Arbeit des Kantons Zürich mit der Zukunft der Künstlichen Intelligenz auseinandergesetzt. Im Rahmen dieser Arbeit wird aufgezeigt, warum lernende Maschinen Teil unseres Alltags werden, welchen Nutzen, aber auch welche Herausforderungen smarte Maschinen für Wirtschaft und Gesellschaft bringen. Und es wird gezeigt, dass auch künftig keine Terminator-Roboter zu erwarten sind, sondern dass die Künstliche Intelligenz fundamental auf die menschliche angewiesen ist.

Im Auftrag von:

W.I.R.E.

WEB FOR INTERDISCIPLINARY RESEARCH AND EXPERTISE

THINK TANK FOR BUSINESS, SCIENCE AND SOCIETY



Kanton Zürich
Volkswirtschaftsdirektion
Amt für Wirtschaft und Arbeit