

Jahrbuch zu den Alpbacher Technologiegesprächen 2018
Alpbach Technology Symposium Yearbook 2018

**Hannes Androsch, Wolfgang Knoll,
Anton Plimon (Hg. Eds.)**

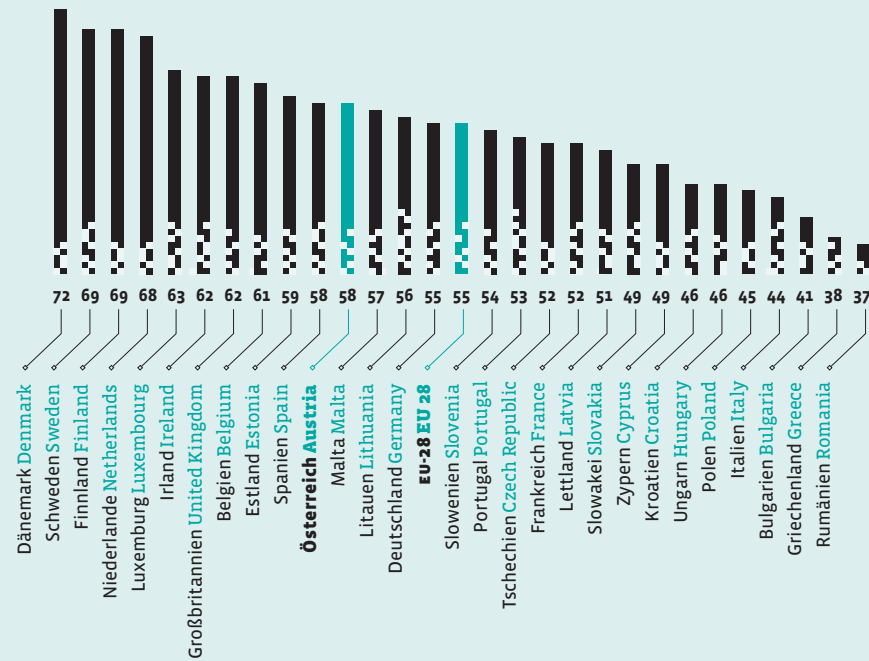
Technologie
im Gespräch
Discussing
Technology

**Künstliche Intelligenz
Artificial Intelligence**

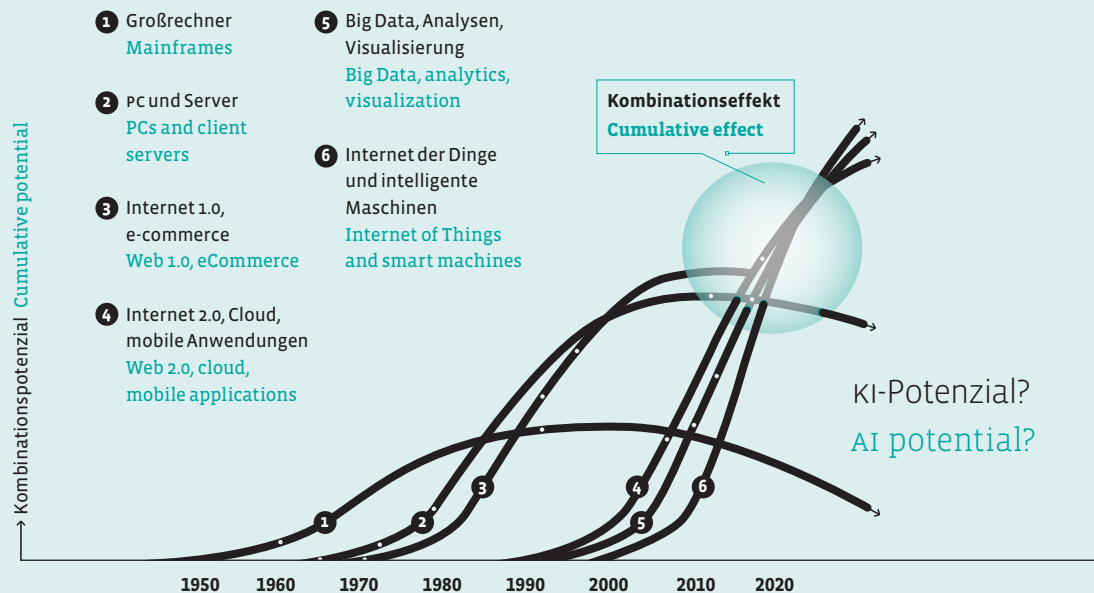
HOLZHAUSEN
— Der Verlag —

Künstliche Intelligenz Artificial Intelligence

1 Wie weit ist die digitale Transformation in Europa fortgeschritten? How advanced is digital transformation in Europe?

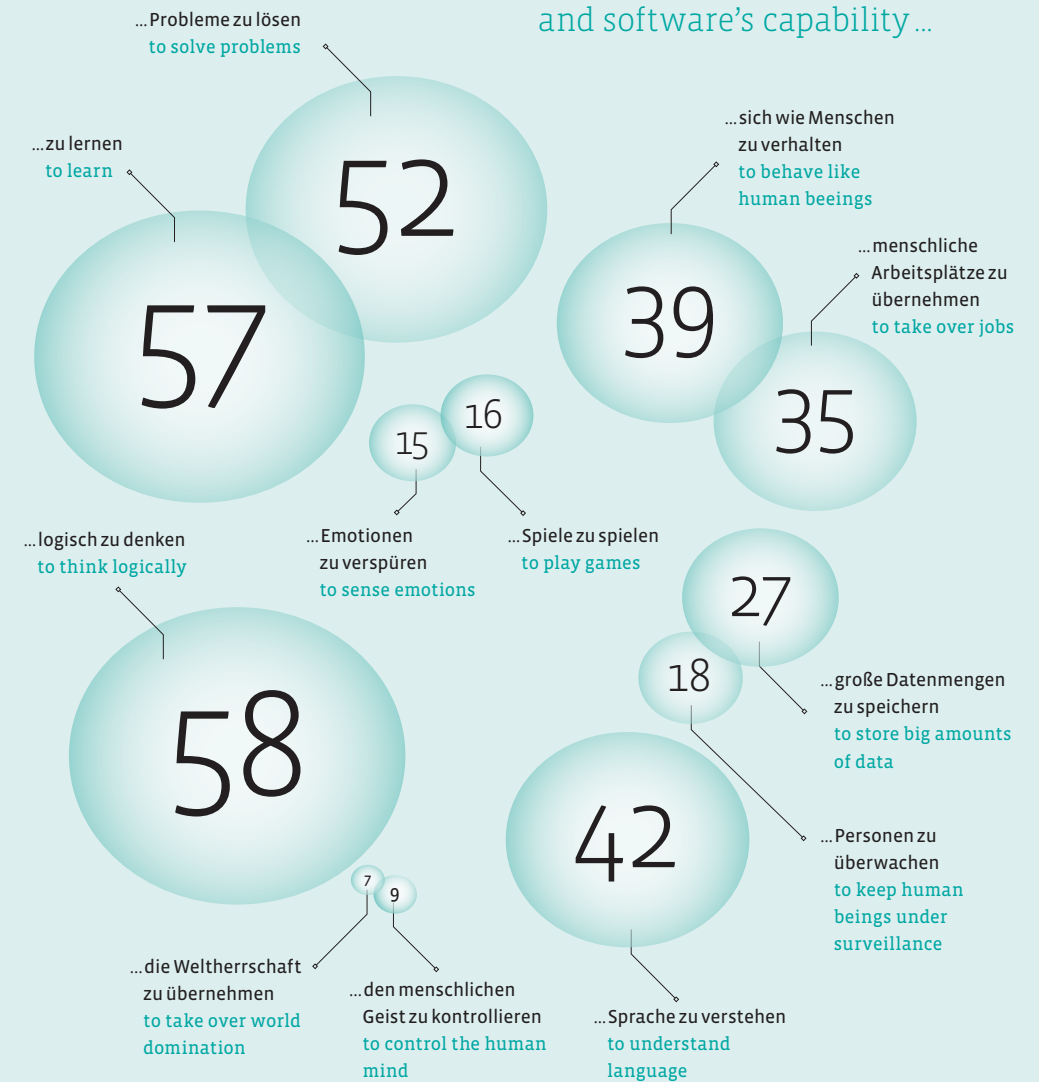


2 Die Kombinationseffekte der Technologien beschleunigen den Wandel. The technologies' combinatorial effects are accelerating change.



3 UMFRAGE: Was verstehen Menschen unter »künstlicher Intelligenz«? SURVEY: What do people mean by the term "artificial intelligence"?

KI ist die Fähigkeit von
Geräten und Software...
AI is the machines'
and software's capability...



Quellen Sources

- European Commission, *Digital Scoreboard*, Stand **Extraction Date**: 2. Juni 2018 [June 2, 2018](#), *Index für die digitale Wirtschaft und Gesellschaft (DES) 2018* [Digital Economy and Society Index \(DESI\) 2018](#).
- <http://reports.weforum.org/digital-transformation/wp-content/blogs.dir/94/mp/files/pages/files/dti-executive-summary-20180510.pdf>; *Digital Transformation Initiative*, 2018 (World Economic Forum/Accenture analysis).
- Pegasystems, *1000 befragte Erwachsene in Deutschland, in Prozent* [1000 responding adults in Germany, in percent](#).

Technologie im Gespräch 2018
Discussing Technology 2018

HOLZHAUSEN
— *Der Verlag* —

Inhalt

Contents

- Hannes Androsch/Martin Kugler
- 6 Bildung und Forschung entscheiden darüber, was wir aus dem digitalen Wandel machen
[Education and Research Are Decisive for What We Make of the Digital Transformation](#)
- Interview Matthias Weber/Petra Schaper-Rinkel
- 18 »Unsere Fähigkeit, sich unser zukünftiges gesellschaftliches Leben vorzustellen, ist begrenzt«
[“Our ability to imagine our social life in the future is limited”](#)
- 36 Was ist Intelligenz, was künstliche Intelligenz?
[What Is Intelligence, and What Is Artificial Intelligence?](#)
- 52 Geschichte der künstlichen Intelligenz
[The History of Artificial Intelligence](#)
- 68 Helmut Leopold
Die Technik hinter der künstlichen Intelligenz
[The Technology Behind Artificial Intelligence](#)
- Interview Manfred Tscheligi
- 80 »Künstliche Intelligenz hat noch keinen Bauch«
[“Artificial intelligence still lacks gut instincts”](#)
- 94 Zwischen Utopie und Dystopie
[Between Utopia and Dystopia](#)
- Helmut Leopold
- 102 Ist KI eine sichere Technologie, und können wir ihr vertrauen?
[Is AI a Safe Technology, and Can We Trust It?](#)
- Interview Allan Hanbury
- 112 »Was man mit Daten tun kann, hat seine Grenzen«
[“There are limits to what can be done with data”](#)
- 128 Die Arbeit der Zukunft oder Die Grenzen des Maschinenlernens
[The Work of the Future or The Limits of Machine Learning](#)

ARTTEC zeigt Schnittstellen von Kunst, Technologie und Wissenschaft:

ARTTEC highlights interfaces between art, technology, and science:



...als Teil dieses Jahrbuches, in einer Ausstellung bei den Alpbacher Technologiegesprächen sowie ab Herbst 2018 im Foyer des AIT-Firmensitzes in Wien. Diesjähriger Kulturpartner ist die Ars Electronica Linz.

...as part of this yearbook, in an exhibition accompanying the Alpbach Technology Symposium, and in the foyer of the AIT headquarters in Vienna from fall 2018. This year's culture partner is Ars Electronica Linz.



138 **Anwendungen Applications** Der digitale Stall: Wie Roboter in unser Leben drängen
[Digital Barns: How Robots Are Pushing into Our Lives](#)

146 **Anwendungen Applications** Georg Langs
Präzisionsmedizin und künstliche Intelligenz
[Precision Medicine and Artificial Intelligence](#)

156 **Anwendungen Applications** Künstliche Intelligenz als dienstbarer Helfer
[Artificial Intelligence as a Subservient Aide](#)

164 **Anwendungen Applications** Helga Nowotny/Stefan Thurner
Wie im Jahr 2038 Entscheidungen gefällt werden
[How Decisions Will Be Made in 2038](#)

174 **Anwendungen Applications** Interview Hermann Hauser
»Wir brauchen Zugang zu den Daten, sonst können wir keine Forschung betreiben«
[“We need access to data, otherwise we cannot do research”](#)

190 **Anwendungen Applications** Gerald Reischl
Digital India: Mit Technologie zurück zur Weltmacht
[Digital India: Technology as a Means to Become a World Power Again](#)

206 **Anwendungen Applications** Alpbacher Technologiegespräche »Diversität und Resilienz«
[Alpbach Technology Symposium “Diversity and Resilience”](#)

»**Die entscheidende Entwicklung der letzten Jahre besteht** aus kultureller Sicht weniger im atemberaubenden Fortschritt des maschinellen Lernens, sondern **in der Tatsache, dass wir uns** aus einem Zeitalter, in dem wir Technologie – also alle Maschinen und Systeme, die wir gebaut haben – als Werkzeuge verwendet haben, **in eine Welt bewegen, in der wir tatsächlich mit Maschinen zusammenzuleben beginnen.**«

Das Zitat stammt aus einer Studie, mit der die Ars Electronica 2017 vom Europarat betraut wurde. Der Europarat ist für die Inhalte der vorliegenden Publikation nicht verantwortlich.

Die abgebildeten Kunstwerke waren Teil des Ars Electronica-Festivals 2017. Auch heuer bietet sich wieder Gelegenheit, Medienkunst vom Feinsten zu erleben: Vom 6. bis 10. September 2018 ruft das Ars Electronica Festival unter dem Titel „Error – The Art of Imperfection“ wieder Hunderte Künstlerinnen und Künstler, Wissenschaftlerinnen und Wissenschaftler, Ingenieurinnen und Ingenieure, Designerinnen und Designer, Technologinnen und Technologen, Entrepreneurinnen und Entrepreneur sowie Sozialaktivistinnen und -aktivisten aus der ganzen Welt nach Linz. Gemeinsam fragen sie nach aktuellen technologischen und gesellschaftlichen Wechselwirkungen und ihren möglichen Ausprägungen in der Zukunft. All das geschieht in und gemeinsam mit einer breiten Öffentlichkeit. Denn Ars Electronica geht aus den klassischen Räumen der Kultur und Wissenschaft hinaus und setzt sich als weltweit einzigartiges Festival für Kunst, Technologie und Gesellschaft mitten in Linz in Szene. Bespielt wird eine Festivalmeile quer durch die Innenstadt mit der POST-CITY, dem stillgelegten Post- und Paketverteilzentrum der österreichischen Post AG, als Festival-Hotspot. Hier finden auch eine Reihe von Festival-Highlights wie die Große Konzernacht mit Chefdirigent Markus Poschner und dem Bruckner Orchester Linz, das große »U19 – CREATE YOUR WORLD«-Festival oder das ebenso dichte wie hochkarätige Konferenzprogramm statt.

→ www.aec.at/error/de/

Gerfried Stocker, Artistic Director, Ars Electronica Linz

The illustrated works of art were part of the Ars Electronica Festival 2017. This year will again offer the opportunity to experience media art at its best: From September 6 to September 10, Ars Electronica is once again summoning hundreds of artists, scientists, engineers, designers, technologists, entrepreneurs, and social activists from all over the world to gather in Linz to jointly investigate current technological and social interdependencies and their potential future manifestations under the title “Error—The Art of Imperfection.” All of this will take place in public and together with people from all walks of life, since a defining aim of Ars Electronica is to take leave of the traditional premises devoted to culture and scholarship and stage an internationally unique festival of art, technology, and society in the very heart of Linz. The locations are to be found along a festival mile across the inner city, with POSTCITY, the former postal service logistics facility, as its hub. It will be the venue for such festival highlights as the Big Concert Night with Markus Poschner conducting the Bruckner Orchestra, the “U19 - CREATE YOUR WORLD” festival, and a jam-packed lineup of conferences featuring stellar participants.

→ www.aec.at/error/en/

“From a cultural point of view, **the decisive development of the last few years is** not so much the breathtaking progress in machine learning but the fact **that** we move from an era in which we used technology—all the machines and systems we built—as tools into a situation in which **we actually begin to live together with machines.**”

This quote comes from a study with which the Council of Europe commissioned Ars Electronica in 2017. The Council of Europe is not responsible for any contents of the present publication.



Hannes Androsch und Martin Kugler

Bildung und Forschung entscheiden darüber, was wir aus dem digitalen Wandel machen

Education and Research Are Decisive for What We Make of the Digital Transformation

Digitalisierung und vor allem künstliche Intelligenz verändern die Welt, unsere Gesellschaft und unser Wirtschaftssystem tiefgreifend. Die neuen Technologien sind freilich in zahlreiche weitere Umwälzungen wie etwa geostrategische oder demografische Verschiebungen eingebettet. Aufhalten lässt sich die Entwicklung nicht, wir müssen uns ihr stellen und die Rahmenbedingungen zukunftsweisend gestalten.

Die Geschichte des Zauberlehrlings, der die Mächte, die er rief, nicht mehr loswurde, gehört zu den großen Mythen der abendländischen Kultur. Johann Wolfgang von Goethe schrieb die Ballade am Vorabend der industriellen Revolution – in einer Zeit also, in der noch nicht absehbar war, dass es tatsächlich einmal Maschinen geben könnte, die autonom handeln und sich sogar der Kontrolle des Menschen entziehen könnten. So etwas gab es damals nur in Märchen und Sagen. Angesichts der Fortschritte bei Digitalisierung und vor allem künstlicher Intelligenz (KI) wirkt der Zauberlehrling heute ziemlich visionär – zumindest wenn man den aufgeregten Diskursen in diversen Medien folgt.

Faktum ist, dass digitale Technologien unser Leben zur Zeit rasant und unumkehrbar verändern. Wir befinden uns am Übergang von der industriellen zu einer digitalen Revolution. Der technologische Wandel ist ein zentraler Teil eines gewaltigen Umbruchs, den die Welt derzeit erlebt. Genauer gesagt: Heute treffen mehrere große Umbrüche zusammen, von denen jeder einzelne schon für sich das Potenzial hätte, die Welt aus den Fugen geraten zu lassen.

Zum einen lässt sich eine völlige Neugewichtung der geopolitischen Situation beobachten. Über allem steht ein geostrategischer Kampf zwischen China und den USA. In diesem Zusammenhang sollte man Russland ebenso wenig vergessen wie das aufstrebende Indien. Europa lässt sich abhängen. Auch Österreich schaffte es in jüngster Zeit nicht, seine Position zu verbessern. Der Kampf wird unter anderem um die Vorherrschaft bei neuen Technologien geführt. Die Wirtschafts- und Machtstrukturen von morgen werden heute gelegt. Die Digitalisierung ist einer, wenn nicht der wichtigste Treiber für Innovationen und Veränderungen von Wirtschaft und Gesellschaft.

Zum anderen erleben wir zwei gegenläufige demografische Entwicklungen: In Indien und in Afrika explodiert die Bevölkerung; das Durchschnittsalter liegt dort bei unter 30 Jahren. Gleichzeitig wird die Gesellschaft in der westlichen Welt immer älter. Es fehlen zunehmend Arbeitskräfte, vor allem junge Fachkräfte, die der Westen für das Funktionieren seines Wirtschafts- und Sozialsystems dringend benötigt. Zur Bewältigung der damit verbundenen Probleme müssen neue Möglichkeiten erschlossen werden, die ebenfalls in die neuen Technologien eingebettet sind.

Digitization and, even more so, artificial intelligence are profoundly transforming the world, our society and economic system. The new technologies come along with other upheavals like geostrategic and demographic shifts. There is no stopping this development; we must face up to it and create viable conditions to lead us into the future.

The story of the sorcerer's apprentice who cannot hold sway over the spirits he has called is one of the great mythological tales of Western culture. Johann Wolfgang von Goethe wrote this ballad on the eve of the Industrial Revolution—at a time, that is, when no one could foresee that some day there would actually be machines able to operate autonomously and even elude human control. At the time, things like this only happened in legends and fairy tales. In view of the progress that digitization and, above all, artificial intelligence (AI) are making, the sorcerer's apprentice appears rather visionary—at least, if one listens in on the frantic buzz of discourses in various media.

The fact of the matter is that digital technologies are about to change our lives rapidly and irreversibly. We find ourselves in the transition from the industrial to a digital revolution. The technological transformation is central to a huge upheaval that the world is currently going through. To be more precise: there are in fact several major upheavals coming together today, each of which would alone have the potential to make the world come apart at the seams.

For one thing, there is a complete realignment of the geopolitical situation that can be observed. Standing above all is a geostrategic struggle between China and the United States. In this context, Russia should not be forgotten, nor should rising India. Europe is falling behind. Austria has not been able either to improve its position in recent times. The struggle is about predominance in new technologies. The foundation of the economic and power structures of tomorrow is laid today. Digitization is one if not the single most important driver of innovation and changes of economy and society.

For another, we are witnessing two contrary demographic developments: in India and Africa, populations are exploding, with the average age being less than thirty years. At the same time, Western societies continue to age. There is an increasing shortage of labor, particularly young skilled labor that the West urgently needs for the functioning of its economic and social system. To cope with the problems entailed, new possibilities must be made accessible, which will also have to be based on the new technologies.

Moreover, digital technologies will also play a central role in dealing with ecological problems like the consequences of climate change or the battle against hunger by facilitating, for example, ecologically sustainable increases of crop production and the fight against air pollution or forest fires.

Reaching all spheres of life Speaking of digitization, most people would probably think of the Internet or robots. However, digitization is a lot more than that. The technologies that will most influence our lives and work already in the near future mainly include, aside from self-driving cars, virtual reality, chatbots, language recognition, and “digital medicine.” These systems are able to learn progressively and respond independently to external conditions. AI and machine learning are being integrated more and more in everyday gadgets and appliances. The Internet of Things and Industry 4.0 are beginning to take shape. In addition, there is an increasing fusion of man and machine: so-called “brain gates”—cerebral implants—have long been tested in humans in research and are widely used in regenerative medicine.

Überdies werden auch zur Bewältigung von ökologischen Problemen, etwa der Folgen des Klimawandels, und im Kampf gegen den Hunger digitale Technologien eine zentrale Rolle spielen, etwa zur ökologisch verträglichen Steigerung von Ernten oder zur Bekämpfung von Luftverschmutzung oder Waldbränden.

Alle Lebensbereiche werden erfasst Bei Digitalisierung denken die meisten wohl an das Internet und an Roboter. Digitalisierung ist aber viel mehr. Zu den Technologien, die unser Leben und unsere Arbeit schon in naher Zukunft am meisten beeinflussen werden, zählen neben selbstfahrenden Autos vor allem Virtual Reality, Chatbots, Spracherkennung und die »digitale Medizin«. Diese Systeme werden zunehmend lernfähig und reagieren selbständig auf externe Gegebenheiten. KI und maschinelles Lernen werden vermehrt in Geräte des täglichen Gebrauchs integriert. Das Internet der Dinge und die Industrie 4.0 nehmen Gestalt an. Hinzu kommt eine zunehmende Verschmelzung von Mensch und Maschine: Schon längst werden in der Forschung sogenannte »brain gates« – Implantate im Gehirn – auch am Menschen getestet und in der regenerativen Medizin vielfach angewendet.

KI könnte von größerer Bedeutung sein als die Zähmung des Feuers oder die Nutzbarmachung der Elektrizität, meinen viele Experten – ähnlich wie vielleicht die Erfindung der Dampfeisenbahn und der Verbrennungskraftmaschine. Die Konsequenzen werden kaum einen Sektor unberührt lassen.

Nehmen wir die Medizin als Beispiel: Unter dem Schlagwort »precision medicine« oder »digital medicine« arbeiten unzählige Forscher und Unternehmen weltweit an neuartigen Behandlungsmethoden für Krankheiten, bei denen jeder Mensch individuell betrachtet wird. Ja mehr noch: Durch eine umfassende Analyse der genetischen Ausstattung und des Lebensstils soll eine wirksame Prävention von Erkrankungen möglich werden. Die Voraussetzung dafür ist, riesige Mengen von Gesundheitsdaten zu erfassen, miteinander zu verknüpfen, Muster und Anomalien zu erkennen und daraus richtige Schlüsse zu ziehen. Man hofft, dadurch beispielsweise Herz-Kreislauf-Erkrankungen oder Krebs schon an der Wurzel packen zu können und erst gar nicht entstehen zu lassen. Im Silicon Valley ist derzeit eine wahre Goldgräberstimmung hinsichtlich der digitalen Medizin ausgebrochen. Große Fonds und Risikokapitalgeber überbieten einander dabei, in die entsprechenden Start-up-Unternehmen zu investieren.

Auch das Bankwesen, um ein zweites Beispiel zu nennen, wird sich völlig verändern: Wir reden heute noch von Bargeld; in Schweden gibt es schon fast keines mehr. Auch die Inder schaffen es gerade ab. Und in China gibt es schon Bettler, die Passanten ein iPhone hinhalten – gespendet wird direkt per Smartphone.

Kurz gesagt: Die Digitalisierung hat mittlerweile alle Lebensbereiche erfasst. Berufsbilder verändern sich, Arbeitsplätze werden wegfallen, neue Jobs werden entstehen. Unternehmen müssen neue Geschäftsmodelle entwickeln und erschließen. Die Technologie wird radikale Innovationen hervorbringen und die Wirtschaft auf eine disruptive Art verändern (Plattformökonomie, Null-Grenzkosten-Gesellschaft, sachwertgeringe Unternehmungen). Mit der Entwicklung sind zahlreiche Probleme und Gefahren verbunden: Man wird heute überall überwacht, egal ob man sein Smartphone eingeschaltet hat oder nicht – das lässt Assoziationen mit George Orwells Big Brother aufkommen. Ein riesengroßes Problem ist die Ausbreitung von Cyberkriminalität, groß sind auch Bedenken hinsichtlich ethischer Konsequenzen und der anfangs erwähnten Zauberlehrlingsauswirkungen, vor denen unter anderem der kürzlich verstorbene Stephen Hawking gewarnt hat.

Nutzen und Gefahr Jede fundamentale Neuerung – zum Beispiel das Messer – hat beide Elemente in sich: Nutzen und Gefahr. Es ist immer eine Sache der Entscheidung, wie man eine Innovation anwendet. Kein Mensch will auf ein Messer verzichten, es kann aber auch zu einer Mordwaffe werden. Das gilt genauso für das selbstfahrende Auto oder für KI. Daher ist es unumgänglich, nicht nur technologische Problemstellungen zu erforschen und zu lösen, sondern sich auch mit rechtlichen, gesellschaft-

Many experts think AI might have greater significance than the taming of fire or the utilization of electricity—similar only to the invention of the steam locomotive and the combustion engine. Its consequences will leave almost no sector unaffected.

Let us take, for example, medicine: under the heading of “precision medicine” or “digital medicine,” countless researchers and enterprises are working on novel treatments for diseases, in which every individual is considered individually. And what is more: comprehensive analyses of genetic setup and lifestyle are supposed to facilitate effective disease prevention. The prerequisite for this is collecting and connecting vast amounts of data, identifying patterns and anomalies, and drawing the right conclusions. It is hoped that it will help to strike at the root of, for example, cardiovascular disease or cancer, or to avert them altogether. A veritable gold rush has descended on “Silicon Valley” with regard to digital medicine. Large funds and venture capitalists are vying for opportunities to invest in start-ups in the field.

To give a second example: banking will undergo a fundamental transformation, too. We still talk about cash today, while in Sweden it almost does not exist any longer. The Indians are also about to abolish it. And there are beggars in the streets of China who hold up iPhones to passers-by—you give directly by smartphone.

In brief: digitization has reached almost all spheres of life. Job profiles are changing, jobs will fall away, new jobs will be created. Corporations will have to find and develop new business models. Technology will bring forth radical innovations and shatter the economy in a disruptive way (platform economy, the zero marginal cost society, low asset-value enterprises). The development, however, brings numerous problems and dangers in its wake: there is mass surveillance everywhere today, whether your smartphone is on or not—which inevitably invokes associations with George Orwell’s Big Brother. One huge problem is rampant cybercrime; also huge are apprehensions with regard to the ethical consequences and the mentioned sorcerer’s apprentice effects, which, among others, also the recently passed Stephen Hawking has cautioned against.

Utility and hazard Every fundamental innovation—for example, the knife—contains in itself both elements: utility and hazard. It is always a matter of making a choice of how to use an innovation. Nobody wants to do without a knife though it may also turn into a murder weapon. The same is true of the self-driving car or AI. It is therefore inevitable not only to explore and solve technological problems but also to deal with the legal, social, judicial, and ethical issues involved. The Facebook scandal which popped up this year in spring—and which in fact only is the tip of the iceberg—clearly shows that the development needs to be based on new rules. These are complex challenges. “Back now, broom, into the closet! Be thou as thou wert before!”—like with Goethe—in any case is a highly unlikely option. And, for us, an “old master” who comes to the rescue of the feckless apprentice is nowhere in sight.

No one has a readymade answer to the challenges brought by the digital upheaval—what we are faced with here is a work in progress. However, solutions are necessary. For progress is coming, whether everybody embraces it or not, whether it instils fears or hopes of bliss and salvation.

lichen, juristischen und ethischen Fragen zu befassen. Der im heurigen Frühling virulent gewordene Facebook-Skandal – der nur die Spitze eines Eisbergs ist – zeigt deutlich, dass die Entwicklung neuer Regeln bedarf. Das sind komplexe Herausforderungen. »In die Ecke, / Besen! Besen! / Seid's gewesen.« – wie bei Goethe – ist eine jedenfalls mehr als unwahrscheinliche Option. Ein »alter Meister«, der dem Zauberlehrling zu Hilfe eilt, steht uns nicht zur Verfügung.

Niemand hat auf die Herausforderungen durch die digitalen Umbrüche eine fertige Antwort – wir stehen vor einem Work in progress. Lösungen sind aber notwendig. Denn der Fortschritt kommt ungeachtet dessen, ob er jedem willkommen ist oder nicht, ob man sich davor fürchtet oder ob man damit Heils- und Erlösungsvorstellungen verbindet.

Wie zahlreiche Beispiele aus der Geschichte zeigen, führt es zu nichts, eine Entwicklung einfach zu verbieten. Ende des 16. Jahrhunderts zum Beispiel hat der englische Student William Lee einen Strumpfwirkstuhl erfunden, der um ein Mehrfaches schneller war als die Herstellung per Hand. Er suchte 1589 bei Königin Elizabeth I. um ein Patent dafür an, das ihm aber verweigert wurde. Offiziell begründete wurde die Ablehnung damit, dass die Qualität der Socken zu grob sei. Man vermutet allerdings, dass das nur eine Ausrede war und die Obrigkeit die Arbeitsplätze der Handstrickerinnen und -stricker schützen wollte. Dieser Verdacht erhärtet sich angesichts der Tatsache, dass Lee 1609 eine verbesserte Maschine anmelden wollte und ihm auch dieses Patent – diesmal von König Jakob I. – abgeschlagen wurde. Aufzuhalten war die Entwicklung trotzdem nicht: Ende des 17. Jahrhunderts gab es mehr als 1500 Wirkstühle in London – und das Königshaus erließ für diese Maschinen ein Exportverbot.

Auch ein Sturm auf Maschinen ist sinnlos, wie etwa der Aufstand der Luddisten zeigt, die zwischen 1811 und 1816 in Nottingham zuerst Nadeln von Strickmaschinen zerbrachen und später ganze Webstühle und sogar Fabriken zerstörten. Diese Gewalt wurde von der Obrigkeit mit Gegengewalt gebrochen – durch 12.000 Soldaten und ein eigenes Gesetz (»Frame-Breaking Act«), das die Zerstörung von Webstühlen unter Todesstrafe stellte. Über ähnliche Turbulenzen berichtet auch Gerhart Hauptmann in *Die Weber*. So nachvollziehbar die sozialen Ursachen für die Aufstände auch sein mögen – aufhalten konnten sie den technologischen Wandel nicht.

Heute will niemand Maschinen tatsächlich zerstören. Eine aktuelle Variante, eine steuerliche Form des Luddismus, ist sicher nicht zielführend. Auch eine solche Antwort wird den technologischen Wandel nicht aufhalten können, zumal nationale Alleingänge ohnehin nicht funktionieren und der Staat zudem seine eigene Politik, nämlich die Förderung neuer Technologien, konterkarieren würde.

Die Zukunft proaktiv gestalten Der digitale Wandel findet statt. Sinnvoll ist es daher, die Veränderungen zu gestalten und geeignete Rahmenbedingungen dafür zu schaffen. Entscheidend ist, wie Gesellschaft, Wirtschaft, Sozialsystem und vor allem Bildungssystem darauf möglichst proaktiv reagieren. Flexibilität wird sowohl aus unternehmerischer als auch aus individueller Sicht notwendig sein, um neue Möglichkeiten nutzen zu können und die Herausforderungen zu bewältigen. Die Schwierigkeit besteht darin, dass man derzeit bestenfalls erahnen kann, welche Entwicklungen eintreten werden, und dass man nicht weiß, wann jene entscheidenden Technologiesprünge tatsächlich stattfinden werden, die erneut grundlegende Veränderungen bewirken werden.

Derzeit sind Europa und Österreich bei digitalen Technologien – und nicht nur in diesem Bereich – im Hintertreffen. Die globalen Akzente werden aktuell in den USA und in China gesetzt. Man darf auch Indien nicht unterschätzen, das mit Riesenschritten aufholt, wie das Beispiel der Technologiemetropole Bangalore zeigt. Europa ist dagegen in der alten Industrie verhaftet: Hier wird über Diesel gestritten, während China zur Elektroauto-Weltmacht wird. Die europäischen Länder müssen hier rasch nachziehen, um im Wettbewerb um datenbasierte Technologien nicht den Anschluss zu verlieren. Es ist daher zu fordern, dass die entsprechenden Mittel für das nächste »Horizon Europe«-Rahmenprogramm der EU für Forschung und Innovation zumindest verdoppelt werden.

As numerous examples from history show, simply trying to prohibit a development does not lead to anything. In the late sixteenth century, the English student William Lee invented a stocking frame that was several times faster than manual production. He applied for a patent with Queen Elizabeth I but was refused it, officially on grounds of the quality of his stockings being too coarse. It was suspected, though, that this was just an excuse and that authorities merely sought to protect hand-knitting jobs—an assumption that was borne out by the fact that when Lee tried to have an improved machine registered in 1609 he was again refused a patent, this time by King James I. The development was nevertheless unstoppable: toward the end of the seventeenth century, there were more than 1,500 stocking frames in operation in London—and the monarchy eventually imposed an export ban on those machines.

Physical attacks against machines do not make sense either, as the riots of the Luddites between 1811 and 1816 show, who first broke off the needles of knitting machines in Nottingham and later took to destroying entire looms and even factories. Their violence was crushed with counter-violence by the authorities—by sending 12,000 soldiers and imposing special legislation (the “Frame Breaking Act”) forbidding destruction of looms under penalty of death. Similar turbulences are recounted by Gerhart Hauptmann in his drama *The Weavers*. However understandable the social causes of such riots may be—they nevertheless could not stop technological change.

Today, nobody actually wants to destroy machines. A present-day variant, a kind of taxation-based neo-Luddism, would certainly not be meaningful. Such a response would not be able to stop technological change either, even less so given that isolated national measures will not work anyway and governments would in effect undermine their own policy of promoting new technologies.

Proactive shaping of the future The digital transformation is happening. It does therefore make sense to try and shape the course of change and create suitable conditions for it. What will be decisive is how society, the economy, the social and, most importantly, the educational system can respond to this as proactively as possible. Both from a corporate and an individual point of view, flexibility will be imperative in order to make use of new possibilities and cope with challenges. The difficulty is that at the moment one can, at best, only guess what developments will become realities, and that nobody knows when those decisive technological leaps will actually occur that will again induce fundamental changes.

Currently, Europe and Austria are lagging behind when it comes to digital technologies—and not only there. Globally, the U.S. and China stake out the terrain. We should not underestimate India, though, which is making giant strides to catch up, as the example of the tech metropolis Bangalore shows. Europe, by contrast, appears caught up in its old industries: here, we argue over Diesel, while China sets out to become a world power of e-mobility. The European countries will have to catch up quickly in order not to be left behind in the competition for data-based technologies. It should therefore be demanded that funding for the next “Horizon Europe” EU framework program for research and innovation should at least be doubled.

Smaller countries like Austria are particularly endangered. This shows, among other things, in this year’s Digital Economy and Society Index (DESI), which is published annually by the European Commission. Like in the years before, Austria is in eleventh place, slightly above the average of the EU-27. Yet the distance to the top, which is dominated by the Scandinavian and Benelux countries, is great. The fact that Germany is shortly behind Austria is no consolation either. We must orient ourselves towards the top, not mediocrity.

Kleinere Länder wie Österreich sind besonders gefährdet. Das zeigt unter anderem der aktuelle Index für die digitale Wirtschaft und Gesellschaft (Digital Economy and Society Index/DESI), den die Europäische Kommission alljährlich veröffentlicht. Österreich kommt wie in den Jahren zuvor auf Platz elf und liegt leicht über dem Durchschnitt der 27 EU-Staaten zu liegen. Doch der Abstand zur Spitze, die von den skandinavischen Ländern und den Benelux-Staaten dominiert wird, ist groß. Dass Deutschland knapp hinter Österreich rangiert, ist kein Trost. Wir müssen uns an der Spitze orientieren, nicht am Mittelmaß.

Der digitale Bereich ist keine Ausnahme, der Rückstand zeigt sich in allen relevanten Rankings. Auch im aktuellen Bericht des Rates für Forschung und Technologieentwicklung zur wissenschaftlichen und technologischen Leistungsfähigkeit Österreichs wird klar nachgewiesen, dass Österreich – im Gegensatz zu den Zielen der 2012 von der Bundesregierung beschlossenen FTI-Strategie – seine internationale Position nicht verbessern konnte. Österreich ist trotz einer Steigerung der F&E-Quote nicht in die Gruppe der »Innovation Leader« vorgestoßen, sondern im Mittelfeld zurückgeblieben. Es gibt ein Missverhältnis zwischen Input und Output: Es gelingt nicht, die wissenschaftlichen Leistungen am Markt umzusetzen und das vorhandene Potenzial auszuschöpfen. Die Gründe dafür sind vielfältig: Das beginnt bei der fehlenden Flexibilität und beim Verwaltungsdschungel und reicht bis hin zu einer zu niedrigen Dotierung der Universitäten und vor allem der Grundlagenforschung. Im aktuellen österreichischen Regierungsprogramm und in den Budgets sind nur wenige entsprechende Maßnahmen zu finden. Es wird zwar eine Digitalisierungsagentur (DIA) eingerichtet, aber es fehlt an Mitteln, den Rückstand an digitaler Infrastruktur abzubauen.

Forschung und Bildung sind entscheidend Eine Schlüsselrolle spielen Bildung und die Qualifizierung von Arbeitskräften. Österreich muss seine jungen Menschen für die Zukunft fit machen, so gut es geht. Derzeit fehlen 60.000 hochqualifizierte Mitarbeiterinnen und Mitarbeiter, und das ist nicht nur den heute niedrigeren Geburtenraten anzulasten. Das bestehende Bildungswesen ist nicht einmal mehr für das Industriezeitalter geeignet, geschweige denn für das digitale Zeitalter. Die Tatsache, dass jeder sechste Pflichtschulabsolvent nicht sinnerfassend lesen kann, zeigt, dass nicht einmal ein ordentlicher Deutschunterricht stattfindet. Es gilt unbedingt danach zu trachten, dass im Bildungssystem niemand abgehängt wird und keine »forgotten people« produziert werden. Wenn man Menschen nicht zukunftsfit macht, werden die schlimmsten Befürchtungen eintreten und die Ängste vor der Digitalisierung zur selbsterfüllenden Prophezeiung.

Die notwendigen Maßnahmen müssen bereits bei der Vorschule ansetzen. Da heute 70 Prozent aller Mütter berufstätig sind, muss schon diese ganztägig sein; für integrationsbedürftige Menschen braucht man erst recht eine Ganztagsbetreuung. Zurzeit gehen viel zu viele Talente verloren – das ist die große gesellschaftliche Herausforderung. Auch die Lehrinhalte müssen sich ändern: Zusätzlich zu profunden Fachkenntnissen müssen den Schülern und Studierenden weitere Qualifikationen vermittelt werden, um für neue Technologien und künftige Herausforderungen bestmöglich gerüstet zu sein. Auf Grundlage eines interdisziplinären Verständnisses gilt es auch gesellschaftliche, rechtliche und ethische Aspekte mit einzubeziehen. Dafür ist das heutige österreichische Schulsystem höchst ungeeignet.

Bildung und Forschung werden darüber entscheiden, was wir aus dem digitalen Wandel machen. Eine Zukunftsvision wäre, dass es möglich sein sollte, in einer besseren Welt zu leben – wobei man natürlich diskutieren kann, was man unter einer besseren Welt versteht. Es geht darum, die Chance zu nutzen. Das ergibt eine Perspektive, und daraus ergeben sich Aufgaben, die es zu bewältigen gilt. Wenn das nicht gelingt, wird die von vielen zu Recht kritisierte Ungleichheit in der Gesellschaft weiter steigen.

The digital area is no exception, the deficit shows in every relevant ranking. The latest report of the Council of Research and Technology Development about Austria's performance in science and technology clearly indicates that—in contrast to the goals of the RTI strategy adopted by the federal government in 2012—Austria could not improve its international position. Despite a stepped-up R&D quota, Austria was unable to make its way up into the group of “Innovation Leaders” but was stuck in mid-table. There is a disproportion of input to output, and we are unsuccessful in taking scientific achievements to the market and exhausting the existing potential. The reasons for this are diverse: from a lack of flexibility and the jungle of public administration to insufficient funding of universities and, above all, fundamental research. Only few measures to address the matter are found in current Austrian government programs and budgets. There are plans to establish a Digitization Agency (DIA), but there still is a shortage of funding to clear the backlog of digital infrastructure.

Research and education are crucial Education and labor qualification play a key role. Austria must try to make its young people as fit for the future as possible. At present, there is a shortage of 60,000 in high-qualification labor, which is not only due to today's lower birth rates. The existing educational system in fact is no longer fit for the industrial, let alone the digital, age. The fact that one out of six of those leaving compulsory secondary school does not have reading comprehension shows that even German language teaching does not work the way it should. It ought to be considered imperative not to leave anybody behind in the educational system and not to produce “forgotten people.” If people are not made fit for the future, the worst apprehensions will come true and the fears of digitization will become a self-fulfilling prophecy.

Necessary measures already have to start in preschool. Given that seventy percent of mothers today are working mothers, all-day school is a necessity, all the more so for people with integration issues. We are currently losing too many talents—that is the main challenge to society. The contents of teaching will have to change as well: apart from in-depth knowledge of special subjects, additional qualifications will have to be imparted to pupils and students to prepare them in the best possible way for new technologies and future challenges. Based on an interdisciplinary understanding, this also involves social, legal, and ethical aspects. This, however, is something that the present-day Austrian school system is highly unsuited for.

Education and research will be decisive for what we make of the digital transformation. One vision for the future could be that it should be possible to make this world a better place—though it is of course debatable what a better world means. It is all about using the chance. That unfolds a perspective, and poses tasks to cope with. If this cannot be done, inequality in our society that is being rightly criticized by many will continue to rise.

Digitization is transforming our world rapidly and irreversibly. We need some new thinking about the distribution of labor, performance, and wealth. In the near future at least, it will still be up to humankind to decide on what objectives to use technologies for and how to specifically implement them. In his “Sorcerer's Apprentice,” Goethe, it seems, already intuited the problems facing us today.

Digitalisierung verändert unsere Welt rapide und unumkehrbar. Wir brauchen ein neues Denken über die Verteilung von Arbeit, Leistung und Wohlstand. Zumindest auf absehbare Zeit ist und bleibt es der Mensch, der darüber entscheidet, für welche Ziele Technologien genutzt und wie sie konkret eingesetzt werden. Goethe hat in seinem »Zauberlehrling« die Problematik von heute offenbar intuitiv gespürt.

Unterschiedliche Sichtweisen auf KI Im Folgenden werden in diesem Jahrbuch zu den Alpbacher Technologiegesprächen aus unterschiedlichsten Perspektiven viele detaillierte Blicke auf das Thema KI geworfen. Um die Konsequenzen von KI für die gesellschaftliche Entwicklung, für das Innovationsgeschehen und für die Wissenschaft dreht sich ein einleitendes Gespräch mit den beiden Innovationsforschern Matthias Weber und Petra Schaper-Rinkel. Diskutiert wird unter anderem, welche Voraussetzungen geschaffen werden müssen, um die Möglichkeiten der Technologie sinnvoll nutzen zu können – sowohl für Mensch und Gesellschaft im Allgemeinen als auch für den Standort Europa und Österreich. Als zentraler Hemmschuh bei solchen Überlegungen erweist sich das menschliche Unvermögen, sich künftige Entwicklungsszenarien vorzustellen.

In einem separaten Block widmen sich Martin Kugler und Helmut Leopold den Grundlagen von KI. Dabei wird zum einen den Begriffen »Intelligenz« und »künstliche Intelligenz« nachgespürt – ist doch alles andere als klar, was genau damit gemeint ist. Um den heutigen Status der Technologie einordnen zu können, wird zum anderen ein historischer Überblick über ein Dreivierteljahrhundert Forschung gegeben. Und schließlich werden die grundlegenden Konzepte, die technischen Voraussetzungen und die aktuellen Methoden von KI näher erläutert.

Ein zentraler Aspekt der Entwicklung von KI ist die Interaktion zwischen Mensch und Maschine. Manfred Tscheligi beklagt im Interview, dass bei der Gestaltung von KI-Systemen derzeit zu wenig auf die Schnittstelle zum Menschen als Nutzer der Technologie geachtet wird. Das berge große Gefahren in sich, die von Unterforderung und »Deskilling« bis hin zu fehlender Akzeptanz der Technologie reichen. Durch überlegtes Design könnten viele dieser Probleme gemildert werden, ist Tscheligi überzeugt.

Utopien und Dystopien Wesen mit Eigenschaften, die gemeinhin dem Menschen zugeschrieben werden, haben immer schon eine große Faszination ausgestrahlt und sowohl positive Utopien als auch schauerhafte Dystopien hervorgerufen. In einem knappen Überblick werden historische Beispiele – etwa der Golem – und die heutige Debatte über KI prägende Schlagworte wie »technologische Singularität«, »Superintelligenz« oder »Transhumanismus« thematisiert.

Viele Befürchtungen hängen auch mit Sicherheitsaspekten der Technologie zusammen. Helmut Leopold fokussiert in seinem zweiten Beitrag auf die Frage, ob KI an sich eine sichere Technologie ist und ob wir ihr vertrauen können. Zum einen ermöglichte die Technologie eine Ausweitung herkömmlicher Formen der Kriminalität, zum anderen würden aber auch völlig neue Formen der Cyberbedrohung auftreten – zu deren Abwehr es erst wenige Methoden gibt, meint Leopold. Manche Befürchtungen erscheinen allerdings angesichts des derzeitigen Entwicklungsstandes von KI deutlich überzogen.

Daten werden heute von vielen Experten als das Gold der Gegenwart und Zukunft angesehen. In einem ausführlichen Interview gibt der jüngst auf eine Stiftungsprofessor an der Technischen Universität Wien berufene Allan Hanbury Einblicke in sein Forschungsgebiet: Data Intelligence. Er erklärt, was man mithilfe von KI alles aus Daten herauslesen kann – und was (noch) nicht, und betont die hohe Bedeutung der Technologiefolgenabschätzung für alle, die in der Grundlagenforschung tätig sind.

Looking at AI from different angles In the following, in-depth looks are taken at the subject of AI from a broad range of perspectives in this yearbook published on the occasion of the Alpbach Technology Symposium. The consequences of AI for the development of society, for innovation, and science is in the center of an opening conversation with innovation researchers Matthias Weber and Petra Schaper-Rinkel. Among other things, they discuss the question of what prerequisites must be provided to make meaningful use of the technology—both for humans and societies in general and for the business location of Europe and Austria. One obstacle that stands in the way of any such considerations turns out to be the human inability of actually imagining scenarios of future development.

In a separate section, Martin Kugler and Helmut Leopold address some of the basics of AI. For one thing, that involves exploring the notions of “intelligence” and “artificial intelligence”—after all, it is anything but clear what exactly they mean. In order to be able to gauge today’s state of technology, a historical survey is provided of three quarters of a century of research. Finally, fundamental concepts, technical prerequisites, and current AI methods are elucidated in some detail.

One central aspect in AI development is the interaction between man and machine. Manfred Tscheligi criticizes in an interview that not enough attention is being given to the interface with the human as the user of the technology. This, he says, involves great dangers, from bore-out and “deskilling” to a lack of acceptance of the technology as such. Tscheligi is convinced that well-thought-out designs could help to alleviate many of these problems.

Utopias and dystopias Creatures showing qualities commonly ascribed to humans have always held great fascination and evoked both positive utopias and horrific dystopias. A brief overview discusses a number of historical examples—such as the golem—and also addresses some of the buzzwords swirling around in today’s AI debate, like “technological singularity,” “superintelligence,” or “transhumanism.”

Many fears are also related to the security and safety aspects of the technology. In his second contribution, Helmut Leopold focuses on the question whether AI is per se a safe technology that we may trust. For one thing, he argues, the technology facilitates the spreading of traditional forms of crime, and for another, it ushers in entirely new types of cyberthreats—with only few methods at hand, Leopold says, to defend against them. Some apprehensions, though, appear to be clearly exaggerated given the current actual state of AI development.

Many experts today consider data to be the gold of the present age and the future. In an in-depth interview, Allan Hanbury, recently appointed to an endowed chair at the Vienna University of Technology, affords an insight into his area of research: data intelligence. He explains what can be read out of data using AI—and what cannot (yet), and he emphasizes the significance of technology assessment for everybody working in basic research.

The role of AI in the work life of the future This also relates to the role that AI is going to have in work life in the future. In a much-noted article they published in *Science* magazine, U.S. researchers Erik Brynjolfsson and Tom Mitchell defined eight criteria to decide in which cases it makes sense to employ AI, and in which not. This leads to several important conclusions that may dampen exaggerated pessimism: for one thing, the usability of machine learning algorithms has its limits, given the current state of development, and for another, only parts of jobs are replaceable by AI.

Die Rolle von KI im Arbeitsleben der Zukunft Das betrifft auch die Rolle, die KI im Arbeitsleben der Zukunft haben wird. In einem aufsehenerregenden Artikel in der Wissenschaftszeitschrift *Science* haben die beiden US-Forscher Erik Brynjolfsson und Tom Mitchell acht Kriterien definiert, in welchen Fällen Methoden der KI sinnvoll einsetzbar sind und in welchen nicht. Daraus ergeben sich mehrere wichtige Schlüsse, die einen übertriebenen Pessimismus dämpfen können: Zum einen sind dem Einsatz von KI beim derzeitigen Entwicklungsstand maschineller Lernalgorithmen Grenzen gesetzt, zum anderen werden immer nur Teile von Jobs durch KI ersetzt werden.

KI-Systeme werden jedenfalls in viele, wenn nicht sogar in alle Bereiche unseres künftigen Lebens Eingang finden. Deutlich sichtbar ist das etwa beim Siegeszug von Robotern – sei es in Fabriken, sei es in der Landwirtschaft, sei es auf der Straße. Wie Georg Langs erläutert, wird KI aber auch in der Medizin als Unterstützung für Ärzte immer wichtiger – etwa um frühzeitig Anzeichen von Krankheiten zu entdecken, Diagnosen zu bewerten oder bei Therapieentscheidungen zu helfen. Helga Nowotny und Stefan Thurner machen einen Blick in eine mögliche Zukunft, wie ein intelligentes Computersystem (das derzeit am Complexity Science Hub in Wien aufgebaut wird) bei Entscheidungen helfen könnte. Und schließlich wird ein knapper Überblick über den aktuellen und künftigen Einsatz von KI in diversen Dienstleistungssektoren gegeben.

Disruption in der Wirtschaft Um die wirtschaftlichen Konsequenzen von KI geht es in einem ausführlichen Interview mit dem IT-Investor und -Vordenker Hermann Hauser. Er sieht bereits einige Bereiche, in denen KI bereits drauf und dran ist, ganze Industriezweige umzukrempeln – etwa durch Spracherkennung oder autonome Fahrzeuge. Den nächsten großen Schritt erwartet er bei der Automatisierung von Entscheidungen. Hauser setzt sich für eine wesentlich klarere Regelung des Datenschutzes ein – denn große anonymisierte Datensätze seien die Grundvoraussetzung für KI, und je besser ein System reguliert ist, desto höher sei die Qualität der Daten.

Den Abschluss dieses Jahrbuches bildet ein Blick in eine der Boomregionen der Welt, nach Indien: Gerald Reischl beschreibt in einer großen Reportage, wie die Regierung mit einer noch nie dagewesenen Digitalisierungsoffensive – mit Big Data, Blockchain, KI und dem Internet der Dinge – die Infrastruktur und die Voraussetzungen schaffen will, um das Land (wieder) zur größten Wirtschaftsmacht und Wissensgesellschaft der Welt zu machen.

Interdisziplinäre Schnittstellen Zwischen den einzelnen Fachbeiträgen finden sich visuelle Beispiele aus dem Bereich der Kunst, deren Urheber sich auf unterschiedliche Weise mit diversen Aspekten von KI auseinandersetzen. Dies ist Teil des Programms ARTTEC, dieses Jahr in Alpbach mit der Ars Electronica realisiert, das Schnittstellen von Kunst, Technologie und Wissenschaft versammelt, um neue sowie zukunftsrelevante Perspektiven zu eröffnen. ✕

Hannes Androsch, geboren 1938 in Wien, ist Aufsichtsratsvorsitzender des Austrian Institute of Technology (AIT), Vorsitzender des Rats für Forschung und Technologieentwicklung (RFTE) und war bis Juni 2016 Aufsichtsratsvorsitzender der Finanzmarktbeteiligungsgesellschaft des Bundes (FIMBAG). In seiner politischen Tätigkeit (SPÖ) war er u. a. Abgeordneter zum Nationalrat (1966–1970), Bundesminister für Finanzen (1970–1981) und Vizekanzler (1976–1981). Danach war er Generaldirektor der Creditanstalt-Bankverein (1981–1988) und Vorsitzender der Österreichischen Kontroll-

bank AG (1985–1986). 1989 gründete er die AIC Androsch International Management Consulting GmbH und begann 1994 den Aufbau einer industriellen Beteiligungsgruppe (Austria Technologie & Systemtechnik AG, Österreichische Salinen AG u. a.). 2004 errichtete er die »Stiftung Hannes Androsch bei der Österreichischen Akademie der Wissenschaften« und ist dort seit 2005 Mitglied des Senats. Ehrendoktorate und Ehrensensator an verschiedenen österreichischen und internationalen Universitäten (Montanuniversität Leoben, Universität New Orleans, USA, u. a.).

In any case, AI systems will be making their way into many if not all fields of our life in the future. This is clearly visible in the triumphant progress that robots are making—whether in manufacturing, agriculture, or on the streets. As Georg Langs explains, AI will also become ever more important in medicine to support doctors—for example, to help early detection of disease symptoms, evaluating diagnoses, and making therapy decisions. Helga Nowotny and Stefan Thurner cast a look into a possible future to indicate how an intelligent computer system (as is currently being built at the Complexity Science Hub in Vienna) may support decision-making. A succinct overview surveys current and future uses of AI in different service industry sectors.

Disruption in the economy The economic consequences of AI are discussed in an extensive interview with IT investor and leading thinker Hermann Hauser. He already sees a number of areas in which AI is set to turn entire industry sectors upside down—for example, through language recognition or self-driving vehicles. The next big step that he expects is automated decision-making. Hauser advocates much clearer data protection rules—large sets of anonymized data are an essential prerequisite of AI, and the better a system is regulated, the higher the quality of the data.

The concluding article of this yearbook takes a look at one of the booming regions of the global economy, India: Gerald Reischl describes in a large reportage how the Indian government is launching an unprecedented digitization offensive—based on big data, the blockchain concept, AI, and the Internet of Things—to establish the infrastructure and the prerequisites needed to make the country (again) the largest economic power and knowledge-based society in the world.

Interdisciplinary interfaces Interspersed between the individual specialized articles are visual examples from the area of art, the creators of which take on different aspects of AI in various ways. This is part of the ARTTEC program, realized this year in Alpbach in cooperation with Ars Electronica, which brings together interfaces of art, technology, and science so as to open up relevant new perspectives of looking to the future. ✕

Hannes Androsch, born in Vienna in 1938, is Chairman of the Supervisory Board of the Austrian Institute of Technology (AIT) and Chairman of the Austrian Council for Research and Technological Development (RFTE). Until June 2016, he was Chairman of the Supervisory Board of the Finanzmarktbeteiligungsgesellschaft des Bundes (FIMBAG). During his political career (SPÖ), his positions included Member of the National Assembly (1966–1970), Federal Minister of Finance (1970–1981), and Vice Chancellor (1976–1981). After this, he served as Director General of Creditanstalt-Bankverein (1981–1988) and as Chairman

of Österreichische Kontrollbank AG (1985–1986). In 1989, he founded AIC Androsch International Management Consulting GmbH, and in 1994 initiated the establishment of an industrial investment group (Austria Technologie & Systemtechnik AG, Österreichische Salinen AG, etc.). In 2004, he founded the “Hannes Androsch Foundation at the Austrian Academy of Sciences,” where he has been a member of the senate since 2005. He has received honorary doctorates and is an honorary senator at various Austrian and international universities, including the Montanuniversität Leoben and the University of New Orleans, USA.

Matthias Weber und Petra Schaper-Rinkel im Gespräch mit Martin Kugler

»Unsere Fähigkeit, sich unser zukünftiges gesellschaftliches Leben vorzustellen, ist begrenzt«



Welche Konsequenzen hat künstliche Intelligenz für Wissenschaft und Innovation? Was bedeutet die Technologie für gesellschaftliche Entwicklungen? Welche Voraussetzungen müssten gegeben sein, damit Europa die Potenziale dieser Technologie ausschöpfen kann? – Ein Gespräch über die Möglichkeiten, mit der rasanten Entwicklung umzugehen.

Welche Konsequenzen hat die Entwicklung von künstlicher Intelligenz (KI) für das Innovationsgeschehen?

Matthias Weber: KI verändert Innovationsprozesse auf zweifache Weise. Die eine hängt mit dem sehr stark ausgeprägten Plattformcharakter zusammen: KI baut auf den Schlüsselkomponenten Daten und Algorithmen auf. Die Plattformen, die beides haben, können schnell wachsen und zeichnen sich durch eine sehr starke Verdrängungskraft aus. Hier gilt häufig das »Winner Takes All«-Prinzip. Diese Plattformen können Standards setzen, auf denen viele weitere Services aufbauen. Rund um sie können sich ganze Ökosysteme von Innovationsaktivitäten entwickeln. In diesen wechselseitigen Abhängigkeiten liegt die sehr starke Pfadabhängigkeit bei der Herausbildung dominierender Plattformen begründet. Der zweite Aspekt ist, dass über KI zumindest Teilaspekte von Innovationsprozessen stärker automatisiert werden können. Invention und Innovation hängen nicht nur mit Kreativität zusammen – es geht auch um die Kombination von Varianten.

Wann werden Inventionen automatisiert?

MW: Zum Beispiel wenn Algorithmen getestet werden – etwa beim automatisierten Fahren. Dabei müssen sehr viele Varianten durchprobiert werden, was über KI automatisiert werden kann. Ein anderes Beispiel sind Algorithmen im Finanzsystem. Auch der Einsatz von Web-Crawling-Tools zur Identifikation von neuen Ideen kann als Teil dieses Automatisierungsprozesses aufgefasst werden. Die Gefahr dieser Entwicklung wird darin gesehen, dass dadurch radikalere Inventionen, die ein hohes Maß an Kreativität verlangen, vernachlässigt werden könnten.

Matthias Weber ist »Head of Center for Innovation Systems and Policy« am AIT. Nach seinem Studium der Verfahrenstechnik und der Politikwissenschaften erlangte er das Doktorat in Volkswirtschaftslehre. In seiner Forschung hat er sich über die vergangenen zwanzig Jahre mit sozio-technischem Wandel in einer Reihe von thematischen Bereichen sowie mit den Veränderungen von Forschungs- und Innovationsmustern auseinandergesetzt. Außerdem berät er nationale Regierungen, europäische Institutionen und internationale Organisationen zu Fragen der Forschungs-, Technologie- und Innovationspolitik. Durch die Kombination von Innovationsystemanalyse und vorausschauenden Ansätzen hat er umfassende Expertise im Bereich »Strategic Foresight« zur Unterstützung der Politik aufgebaut. Als Mitglied mehrerer hochrangiger Gruppen von Expertinnen und Experten und als Koordinator europäischer Forschungsprojekte hat er jüngst den europäischen Forschungskommissar beraten und die Vorbereitung des nächsten Rahmenprogramms für Forschung und Innovation unterstützt.

Matthias Weber is Head of Center for Innovation Systems and Policy at AIT. He has a background in process engineering and political sciences and holds a doctorate in economics. For more than twenty years, he has been doing research on socio-technical change and new patterns of research and innovation in a wide range of thematic domains and advising national governments, European institutions, and international organizations on matters of research, technology, and innovation policy. By combining innovation systems analysis with forward-looking approaches, he has built up extensive expertise on matters of strategic foresight in support of policy-making. As a member of several high-level expert groups and leader of research and consultancy projects, he has recently been advising the European research commissioner and supporting the preparation of the next framework program for research and innovation.

Matthias Weber und Petra Schaper-Rinkel in an interview with Martin Kugler

“Our ability to imagine our social life in the future is limited”

What are the consequences of artificial intelligence for science and innovation? What does the technology mean for social developments? What would the prerequisites be for Europe to make full use of the potentials of this technology?—A conversation about possible ways of dealing with this rapid development.

What are the consequences of the development of artificial intelligence (AI) for what’s happening in innovation?

Matthias Weber: AI transforms innovation processes in two ways. One of them has to do with its distinct platform character: AI builds on data and algorithms as key components. Platforms that have both can grow quickly and have the strength to squeeze out others. “The winner takes it all” is the principle that often applies here. These platforms can define standards that many other services are based upon. Around them, entire ecosystems of innovation activity can develop. These interdependences also are the reason for the strong path dependence in the emergence of dominant platforms. The second aspect is that at least some aspects of innovation processes may be more effectively automated through AI. Invention and innovation are not only related with creativity—they are also about how to combine variants.

Can you give an example of automated invention?

MW: For example, when algorithms are being tested—like in self-driving cars. This involves trying out a large number of variants, which can be automated through AI. Another example are algorithms in the financial system. The use of web-crawling tools to identify new ideas may also be understood as part of this automation process. The danger of this development is seen in the fact that this may lead to disregarding more radical inventions that require a high amount of creativity.


Petra Schaper-Rinkel: Ein weiteres Beispiel ist die Schnittstelle von KI und Biotechnologie bzw. synthetischer Biologie. Da mithilfe von KI viele genetische Variationen simuliert werden können, sind Testreihen mit einer unvorstellbar hohen Geschwindigkeit möglich. Wir gehen im Moment sehr von KI als einer Annäherung an menschliche Intelligenz aus. Weniger im Blick haben wir, dass KI auch ganz eigene Innovationen in Angriff nehmen kann, die fern unserer Innovationsstätigkeit verlaufen. KI kann mit uns als Objekten Innovationsmuster generieren, die wir uns nicht vorstellen können und die weit außerhalb dessen liegen, was wir als intentionales Handeln begreifen. So kann KI etwa Testreihen mit dem Verhalten unterschiedlicher Nutzergruppen ausführen. Google oder Amazon können Geschäftsmodelle testen, von denen wir nicht einmal wissen, dass sie diese testen. Die Begrenztheit von KI wird immer wieder daran festgemacht, dass kein intentionales Handeln und keine Emotion vorhanden sind. Wir haben noch keinen Begriff davon, was es heißt, dass die Logik von KI nicht der Logik traditioneller politischer Prozesse unterworfen ist, und verstehen nicht, dass das kein Defizit von KI ist, sondern in bedrohlicher Weise gesellschaftlich bestimmend wird.



Dr. Petra Schaper-Rinkel, Politikwissenschaftlerin und Senior Scientist am AIT, forscht zu Dimensionen der technischen Zukunft und zu den Instrumenten, mit denen sich Zukunftstechnologien gesellschaftlich gestalten lassen. In EU-Forschungsprojekten entwickelt sie partizipative Zukunftsprozesse, -szenarien und Metanarrative und analysiert Strategien der Zukunftsgestaltung und die politischen Erwartungen an Zukunftstechnologien. Sie koordiniert die inhaltliche Ausrichtung und Programmentwicklung des Forschungsbereichs »Social Futures« und ist am AIT für Technikfolgenabschätzung für das österreichische Parlament verantwortlich.

Sie führen viele Foresight-Prozesse durch: Was kann man daraus lernen?

MW: Zum einen gibt es die Strömung der sogenannten Futuristen, die sich mit grundsätzlichen philosophischen Themen auseinandersetzen und sich fragen, ob KI die menschliche Intelligenz ersetzen wird und was das für die Zukunft der Menschheit bedeuten könnte. Daneben gibt es eine ganze Reihe von Arbeiten, die sich damit beschäftigen, wie KI in spezifische Systeme eingebettet werden kann – z. B. beim autonomen Fahren, in Finanzmärkte, in Energiesysteme, in den Handel usw. Darüber hinaus gibt es allerdings nur sehr wenige Betrachtungen über die weiterführende Rolle von KI in unserer Gesellschaft, etwa im Hinblick auf Abhängigkeiten, die sich aus dem extensiven Einsatz von KI ergeben, oder darauf, welche neuen Chancen und Risiken KI bei demokratischen Entscheidungsprozessen eröffnen kann. Dieser Mangel scheint darin begründet zu sein, dass es uns an Kreativität und Vorstellungskraft mangelt, solche Überlegungen anzustellen. Ilkka Tuomi hat den Begriff der »ontologischen Expansion« verwendet. Der Begriff besagt, dass unsere Fähigkeit, sich unser zukünftiges gesellschaftliches Leben vorzustellen, begrenzt ist und sich diese Vorstellung nur schrittweise mit dem Entstehen neuer Systeme erschließt. Das Bild davon, welche Möglichkeiten KI bietet, erweitert sich also kontinuierlich.



Dr Petra Schaper Rinkel, political scientist and Senior Scientist at AIR's Center for Innovation Systems & Policy, researches on dimensions of the technological future and instruments for the societal design of tomorrow's technologies. In EU research projects, she develops participatory future processes, scenarios, and meta-narratives and analyses strategies of designing the future and political expectations concerning future technologies. She coordinates the orientation and program development of the research field "Societal Futures" and is the AIR's scientist responsible for technology assessment and foresight for the Austrian Parliament.

Petra Schaper-Rinkel: Yet another example is the interface of AI and biotechnology, or synthetic biology. The fact that many genetic variations may be simulated with the help of AI makes incredibly high-speed test series possible. At the moment, we prefer to think of AI as an approximation to human intelligence. What we have less in mind is that AI may also start its own innovations occurring far outside the scope of our innovation activities. AI may use us as objects to generate innovation patterns that we cannot even imagine and which are far beyond anything we would define as intentional action. For example, AI could conduct test series on the behavior of different groups of users. Google or Amazon might be testing business models without us even knowing as much as that they are being tested. The limitations of AI are always said to be the lack of intentional action and emotion. We have no idea yet what it means that AI logics are not subject to the logic of traditional political processes, and we don't understand that this is not a deficiency of AI but something that is about to become predominant in society in a threatening way.

You conduct many foresight processes: what can be learnt from them?

MW: For one thing, there is the school of the so-called futurists who address fundamental philosophical issues, asking themselves whether AI will ever replace human intelligence and what this might mean for the future of humankind. Aside from that, there is a number of projects that deal with how to embed AI in specific systems—for example, in self-driving cars, financial markets, energy systems, trade, etc. Beyond that, however, there are only very few considerations of what further role AI might have in our society, like with respect to dependencies resulting from the extensive use of AI, or what new chances and risks AI may bring to democratic decision-making processes. This lack appears to be rooted in that we lack the creativity and imagination to make such considerations. Ilkka Tuomi has used the term "ontological expansion," meaning that our ability to imagine our social life in the future is limited, and that any such imagination only unfolds gradually with the emergence of new systems. The picture of what possibilities AI offers thus is continually widening.

PSR: Wir können uns gerade einmal vorstellen, was in Bereichen geschieht, die wir überblicken können. Aber von der Ebene darüber können wir uns kein Bild machen. Ende der 1990er-Jahre wurde in Delphi-Studien noch prognostiziert, dass heute alle neu gebauten Häuser einen Medienraum haben würden. Damals dachten auch führende Expertinnen und Experten nicht daran, dass viele Menschen einmal die digitale Technik in Form von Handys und Tablets immer mit sich führen würden.

MW: Bevor Twitter erfunden wurde, konnte man sich nicht vorstellen, dass Präsident Obama in einer öffentlichen Veranstaltung eine Twitter-Meldung senden würde oder dass jemand wie Präsident Trump das Medium zum ultimativen Governance-Instrument entwickeln würde. Erst die Kombination von Internet, Mobilkommunikation und der speziellen Innovation eines Kurznachrichtendienstes hat dieses gesellschaftliche Phänomen möglich gemacht. Das ist ein typisches Beispiel für eine »ontologische Expansion«.

PSR: Wir stellen uns viel zu wenig die Frage, welche Entwicklungsoptionen KI selbst hat. KI entwickelt sich anhand von Zieldimensionen. Die Zieldimension einer KI von Google und Facebook ist auf die Optimierung von Geschäftsmodellen gerichtet. Ist eine KI denkbar, die tatsächlich individuell kontrollierbar und gestaltbar ist? Eine KI, die gesellschaftspolitischen Prioritäten folgt, die zum Beispiel auf Nachhaltigkeit und Gerechtigkeit setzt? Welche Voraussetzungen bräuchten wir dafür? Auf jeden Fall wäre das nicht die KI, die Google oder Facebook im Moment weiterentwickeln.

Die Steuerung der Gesellschaft und politischer Systeme war immer schon schwierig. Wird das durch Technologien wie KI noch schwieriger?

MW: Es gibt Optimisten, die sagen: KI-Systeme erlauben es, komplexe Phänomene, mit denen die Gesellschaft konfrontiert ist, gründlicher zu erfassen, zu simulieren und darauf aufbauend Strategien zu entwickeln; der Klimawandel ist ein typisches Thema. Pessimistische Stimmen argumentieren hingegen, dass KI- und vor allem Big-Data-Analysen auf Mustern der Vergangenheit aufbauen und davon ausgehen, dass sich diese in Zukunft fortsetzen werden. Das mag in vielen Fällen plausibel erscheinen, muss aber nicht notwendigerweise so sein. Damit verbessert sich zwar die Fähigkeit, komplexere Zusammenhänge zu erfassen und eventuell sogar in Echtzeit nachzujustieren und steuernd einzugreifen, speziell bei technischen Systemen. Aber das ändert grundsätzlich nichts an der Fähigkeit einer Gesellschaft, mit

PSR: We can barely imagine what is going on in areas that we are able to overlook. But there is no way that we can get a picture of the level beyond that. At the end of the 1990s, Delphi studies still predicted that all new-built houses would have a media room today. Back then, even leading experts in the field could not imagine that people would carry digital technology on them as smartphones or computer tablets one day.

MW: Before Twitter was invented nobody would have imagined that president Obama would ever send out a tweet from a public event or that somebody like president Trump would develop the medium into the ultimate instrument of governance. It took the combination of Internet, mobile communication, and the specific innovation of a short messaging service to make this social phenomenon possible. This is a typical example of “ontological expansion.”

PSR: The question we don’t ask enough is what developmental options AI itself has. AI develops in terms of goal dimensions. The goal dimension of AI by Google or Facebook is focused on optimizing business models. Is it conceivable to have AI that actually is individually controllable and configurable? AI that follows sociopolitical priorities and, for example, goes for sustainability and justice? What would be the prerequisites for this? In any case, it would not be the kind of AI that is being developed by Google or Facebook right now.

Social and political steering has always been difficult. Do technologies like AI make it even more difficult?

MW: There are optimists who say that AI systems permit to explore in-depth, and simulate, complex phenomena that society is confronted with, and to develop strategies on this basis; climate change is one typical issue. Pessimistic voices rather argue that AI and big-data analyses rely on past patterns, assuming that they are going to continue into the future. That may seem plausible in many cases but is not necessarily so. It improves the ability of understanding more complex connections and perhaps even readjusting them or intervening in them in real time, particularly in technical systems. But, basically, it doesn’t change anything about a society’s ability of dealing with radical and disruptive change for which there are no patterns from the past. One radically skeptical take on AI is based on the conviction that the whole thinking on the issue is too simple, and that the

radikalen und disruptiven Veränderungen umzugehen, für die es keine Muster aus der Vergangenheit gibt. Ein radikal skeptischer Zugang zu KI ist von der Überzeugung getragen, dass das alles viel zu einfach gedacht ist und die wirklich entscheidende Frage lautet, wie in Zukunft Entscheidungsprozesse organisiert werden sollen, in denen KI zwar eine wichtigere Rolle spielen soll, bei denen es aber nicht zu einem Verlust an autonomer Entscheidungsfähigkeit des Einzelnen oder gar einer Entmündigung kommen darf. Wohin das führen kann, sehen wir gerade an den ersten Erfahrungen, die in China mit seinem »Sozialkreditsystem« gemacht werden. Es geht also um die Frage, wie weit man sich bei Entscheidungen – seien sie individuell oder gesellschaftlich – auf KI verlassen will.

PSR: Entscheidend ist auch, wie die extreme Geschwindigkeit des Maschinenlernens mit dem notwendigen Zeitbedarf einer demokratischen politischen Öffentlichkeit kollidiert, die des gemeinsamen Handelns und Sprechens bedarf: Wie schnell kann staatliche Politik in Zukunft noch steuernd eingreifen, wenn Programme führender Unternehmen in der Lage sind, jede politische Maßnahme, die gerade diskutiert wird, schon auf ihre Auswirkungen hin zu simulieren und im Zweifelsfall rasch Gegenmaßnahmen zu entwerfen? Wie stark kann politische Kontrolle zum Beispiel über neue Formen der Besteuerung internationaler Konzerne überhaupt noch wirksam werden?

Manche Experten sind der Meinung, dass KI ein Werkzeug sein kann, das Menschen ermächtigt, selbst eine »bessere«, zutiefst demokratische Gesellschaft zu organisieren.

PSR: Einem sehr optimistischen KI-Szenario zufolge sollte jedem von uns die gleiche avancierte KI zur Verfügung stehen, die nicht in die Systeme der großen Player eingebaut ist. Dann könnte jeder zu einem optimalen Marktteilnehmer werden und selbst über alles entscheiden. Diese Vision ist aber unrealistisch, denn KI braucht viele Daten – und die haben nur die großen Player. Doch selbst wenn die Realisierung dieses Szenarios möglich wäre, hätte man damit nichts gewonnen: Denn dann wären wir völlig atomisiert; jede/jeder Einzelne würde nur noch ihren/seinen Vorteil suchen. Die Technologie würde den individualisierten Konsumentinnen und Konsumenten, dem *Homo oeconomicus*, machtvolle Instrumente liefern und das Leben der Bürgerinnen und Bürger, das *Zoon politikon*, bis zur Abschaffung in den Hintergrund drängen. Persönliche Assistentinnen und Assistenten würden den billigsten Kredit, den günstigsten

real crucial question is how to organize decision-making processes, in which AI can play a more important role but which do not entail a loss of the individual's personal autonomy of decision, or even their incapacitation. Where that may lead can be seen from the first experiences that China is making with its "social credit system." So this is about the question in how far we want to rely on AI in making decisions, individually and socially.

PSR: It is also crucial to see how the extreme speed of machine learning collides with the necessary time needed by a democratic political public that depends on communal speaking and acting: how quickly will government policies be able to intervene in the future when programs operated by leading companies are able to simulate the effects of any political measure that is being discussed and, in case of doubt, come up with swift counteraction? How will political control, for example, by imposing new types of taxation on big international corporations, be able to have any effect at all?

Some experts are convinced that AI can be a tool to empower people to organize a "better," profoundly democratic society for themselves.

PSR: According to a very optimistic AI scenario, each of us should have the same advanced AI available to them, which is not built into the systems of the big players. Then everybody could become an optimal market participant and decide everything for themselves. However, this vision is unrealistic because AI needs a lot of data—and these are in the hands of the big players. Yet even if this scenario could be made a reality, nothing would be gained: for it would mean complete atomization, with each and every individual only seeking their personal advantage. Technology would provide individualized consumers, the homo economicus, with powerful instruments and force their life as citizens, as political animals, into the background, almost to the point of extinction. Personal assistants would find the cheapest loan, the most affordable trades business, the most advantageous terms and conditions and the political party that best represents their special interests. Everybody would be even more easily susceptible to manipulation. Society and democracy would vanish because in this AI development path of optimizing market partici-

Handwerksbetrieb, die vorteilhaftesten Geschäftsbeziehungen und die Partei suchen, die Partikularinteressen am besten vertritt. Dann wären alle noch viel stärker durch Manipulationen angreifbar. Gesellschaft und Demokratie würden verschwinden, weil es in diesem KI-Entwicklungspfad von optimierenden Marktteilnehmern keine kollektiven Prozesse mehr gibt. Wir brauchen Handlungsmacht und Orte, die gesellschaftliche und politische Steuerung im Sinn einer demokratischen Kontrolle überhaupt möglich machen.

MW: Steuerung hat immer auch etwas mit einem gemeinsamen Bezugsrahmen zu tun, in dem gesteuert wird und der sich etwa auch auf gemeinsame Wertvorstellungen bezieht. Bei einer atomisierten Gesellschaft gibt es keine Bezugsrahmen dieser Art mehr, außer vielleicht jene, die in den Echokammern der sozialen Medien entstehen. Das zeigt, dass man grundsätzlich auch innerhalb der virtuellen Welt Bezugsrahmen schaffen kann. Bislang ist es jedoch nur selten gelungen, diese zu einem Forum breiteren öffentlichen Diskurses zu machen, in dem kontrovers und zugleich sachlich argumentiert werden kann.

Welche Veränderungen bringt KI für die Wissenschaft?

PSR: Wir sind es gewohnt, dass die Grundlagenforschung an öffentlichen Einrichtungen und Universitäten geschieht und dadurch eine breite Zugänglichkeit des Wissens gewährleistet ist. KI wandert aber in den privatwirtschaftlichen Raum. Das hat nicht nur etwas mit Geld zu tun: Wenn europäische Spitzenuniversitäten auch in der Lage wären, den führenden KI-Forscherinnen und -Forschern höhere Gehälter zu zahlen als Google und Facebook, wäre für deren Praxis und Theorie Google und Facebook einfach attraktiver. Denn dort sind die Daten, Algorithmen und Infrastrukturen vorhanden, mit denen sie experimentieren können. Wird man die Spitzenforschung im Bereich der KI überhaupt an Universitäten und öffentlichen Einrichtungen halten können – oder sind nicht die Infrastruktur der Forschung und die Daten bereits so sehr privatisiert, dass Spitzenforscherinnen und -forscher zu Unternehmen gehen müssen? Wie sollen wir überhaupt noch beurteilen, an welchem Punkt wir stehen?

pants there is no more room for collective processes. We need power to act and places that facilitate social and political steering in the sense of democratic control in the first place.

MW: Steering also always has to do with a common frame of reference within which the steering is done and which also relates to shared ideas and values. In an atomized society, there are no frames of reference any more, except perhaps for those unfolding in the echo chambers of social media. This shows that, basically, frames of reference may also be established inside the virtual world. So far, though, they rarely could be successfully turned into a forum of broader public discourse, a site of controversial and yet factual argument.

What changes is AI going to bring for science?

PSR: We are used to fundamental research being conducted at public institutions and universities, warranting broad accessibility of knowledge. However, AI is moving into the sphere of private enterprise. This does not only have to do with money. Even if European top universities were able to pay higher salaries to leading AI researchers than they are offered from Google and Facebook, Google and Facebook would still be more attractive for their practice and theory. Because this is where the data, algorithms, and infrastructures are that they need to experiment with. Will it be at all possible to keep cutting-edge research in the area of AI at universities and public institutions—or haven't research infrastructure and data been privatized to such an extent that top researchers simply have to move over to the private sector? How are we supposed to still be able to judge what point we are at?

MW: The question is what scope this development leaves for our universities. They may still have a function in education and training, and in some areas of fundamental research such as algorithm development. As soon, though, as it's about applications, things must occur in the private sector because it is only there that central elements like the data needed are available.

MW: Die Frage ist, welchen Handlungsspielraum diese Entwicklung für unsere Universitäten noch übrig lässt. Diese können weiterhin eine Funktion in der Ausbildung und in manchen Bereichen der Grundlagenforschung, etwa in der Algorithmen-Entwicklung, übernehmen. Sobald es jedoch um Anwendungen geht, muss das im privatwirtschaftlichen Raum geschehen, weil nur dort so zentrale Elemente wie die erforderlichen Daten zur Verfügung stehen.

PSR: In der Innovationsforschung gibt es den Begriff der »Technoscience«. Das bedeutet, dass Forschung und Forschungsgegenstand einander bedingen. Das ist bei KI exemplarisch gegeben: Der Algorithmus, der als Grundlage für die Forschung dient, ist – wenn er funktioniert – sofort das Produkt. Jede Investition in die Grundlagenforschung der Algorithmen kann unmittelbar zu neuen Produkten führen.

MW: Früher hatte die Grundlagenforschung eine größere Distanz zur Anwendung, und diese Anwendungsferne ist durchaus zu Recht kritisiert worden. Das ist heute ganz anders: Heute gibt es ein sehr starkes Wechselspiel zwischen Grundlagenforschung und neuen Produkten oder Dienstleistungen. Angesichts dieses Sogs zur Anwendung stellt sich die Frage, ob die experimentellen Freiräume der Grundlagenforschung vor diesem Hintergrund in Zukunft nicht wieder ausgebaut werden sollten.

Was bedeutet all das für die Forschungs-, Technologie- und Innovationspolitik?

MW: Die Geschwindigkeit der Innovationsprozesse nimmt stark zu. Darauf kann die herkömmliche FTI-Politik mit ihrer eher langsamen Herangehensweise nicht mehr schnell genug reagieren. Zudem stellt sich die Frage, mit welchen Hebeln FTI-Politik überhaupt eingreifen kann. Die staatliche Forschungsförderung wird von der Finanzstärke globaler Unternehmen bei Weitem ausgestochen. Entscheidungen, die im Silicon Valley getroffen werden, sind viel gestaltungsmächtiger als nationale FTI-Politik. Das bedeutet natürlich nicht, dass die klassischen Steuerungsinstrumente der FTI-Politik irrelevant werden, aber im Verhältnis dazu werden andere Fragestellungen wichtiger werden, die sich auf die weitreichenden Auswirkungen von KI und die Anforderungen für deren sinnvollen Einsatz beziehen. FTI-Politik wird dadurch de facto Gesellschaftspolitik. Da geht es dann um Fragen wie: Wie soll denn eine KI-basierte Gesellschaft aussehen? Wie viel Transparenz, wie viele proprietäre Sys-

PSR: In innovation research, there is the term “techno-science.” It means that research and the research subject are mutually conditional. AI is exemplary of this: the algorithm which provides the basis of research is—if it works—at once a new product. Any investment in fundamental algorithm research may immediately lead to new products. Given this gravitational pull toward application, the question arises whether the experimental leeway of fundamental research should not be expanded again in the future.

What does all that mean for research, technology, and innovation policy?

MW: The speed of innovation processes is picking up. This is something that traditional RTI policies with their rather slow approach can no longer respond to quickly enough. Also, the question arises what leverage RTI policy-makers have in order to intervene. Government research promotion and funding are by far outweighed by the financial strength of global corporations. Decisions that are made in Silicon Valley have a lot more impact than national RTI policies. Of course, this does not mean that the classical steering instruments of RTI policy are becoming altogether irrelevant, but by comparison, other issues relating to the far-reaching effects of AI and the prerequisites of its meaningful employment are becoming more important. RTI policy thus de facto becomes social policy. It will then be about issues like the following: What is an AI-based society supposed to look like? How much transparency, how many proprietary systems should there be? Or shouldn't those systems rather be open source? Here, for example, legal frameworks play a decisive role. We must learn to better understand the context in which AI is used and to create suitable conditions that help avoid dependencies and reduce negative effects.

That means that there must be a much broader social debate on how to shape our future...

PSR: Decisions are not primarily made in democratically legitimized contexts today but by the big players in technology. They are making the policies that shape global society. Our current social policies only see the impact of technologies on labor, like the emergence of the crowd-working sector. Nobody thinks about taking care of the design of the infrastructures that, in terms of labor and social policy, are central for the future. In the nineteenth century, telecommunication or the railway

teme soll es geben? Oder sollen die Systeme nicht besser offen sein? Hier spielen etwa rechtliche Rahmenbedingungen eine entscheidende Rolle. Wir müssen den Kontext, in dem KI eingesetzt wird, besser verstehen lernen und geeignete Rahmenbedingungen schaffen, die Abhängigkeiten vermeiden und negative Auswirkungen reduzieren helfen.

Das heißt, dass es eine viel breitere gesellschaftliche Diskussion über die Gestaltung der Zukunft geben müsste...

PSR: Die Entscheidungen werden heute nicht primär in demokratisch legitimierten Kontexten, sondern von den großen technologischen Playern getroffen: Sie machen weltweit Gesellschaftspolitik. Die derzeitige Sozialpolitik sieht nur die Auswirkungen der Technologien auf die Arbeit wie etwa das Entstehen des Crowdfunding-Bereichs. Aber es kommt niemand auf die Idee, sich um die Gestaltung der Infrastrukturen zu kümmern, die arbeits- und sozialpolitisch zentral für die Zukunft sind. Im 19. Jahrhundert hat man aus guten Gründen die Telekommunikation oder die Bahn quasi nationalstaatlich organisiert. Man hat damals gesagt: Das sind die Infrastrukturen, auf denen der Fortschritt beruht. Heute kann man diese Infrastrukturen nicht mehr staatlich aufbauen – das wäre auch Unsinn –, aber sie müssten so gestaltet werden, dass sie für alle Akteure gleichermaßen zugänglich sind.

Was soll man sich darunter vorstellen?

PSR: Wenn etwa der Handel in Europa auf einer kollaborativen Plattform stattfinden würde statt auf Amazon, ließe sich der Zugang zu den Daten demokratisch regeln und die Daten – beispielsweise die Veränderungen der Nutzerprioritäten – könnten in jeweils unterschiedlicher Anonymisierung den Konsumentinnen und Konsumenten sowie Firmen zur Verfügung stehen. Dann gäbe es Transparenz und die Möglichkeit der Steuerung, dann könnten auch gesamtgesellschaftliche Vorteile des Wohlfahrtsstaats allen – den Bürgerinnen und Bürgern, den Konsumentinnen und Konsumenten, den Unternehmen, der Industrie – zugutekommen.

system were organized in a quasi-nationalized way. People said back then: these are the infrastructures that progress is based upon. Today, such infrastructures can no longer be built by states—which would not make sense anyway—but they have to be designed in such a way as to be equally accessible to all actors.

How does one have to imagine this?

PSR: If, for example, retail trade in Europe took place on a collaborative platform instead of Amazon, data access could be democratically regulated, and the data—for example, about changes of consumers priorities—could, with different degrees of anonymization, be made available to consumers and businesses alike. Thus, there would be transparency and the possibility of steering things, and the gains of the welfare state could be afforded to everybody—citizens, consumers, companies, the industry.

MW: From a nation-state perspective, this also enables creating possibilities of embedding AI in different economic activities expeditiously enough. This is an important pillar of RTI policy, which will continue to exist. As said before: the sociopolitical dimension of RTI policies will even be strengthened in the years to come.

However, the sociopolitical discussion that would be necessary is currently not being held. Where could it take place, and how could it be cranked up?

WM: Such discussions could take place in the sphere of national parliaments, which, after all, are supposed to be a forum of public debates in which, in representation of the citizenry at large, different political views and ideas for the future are articulated. The German Bundestag, for example, has been conducting debates on controversial technological subjects for many years; there are so-called Enquete-Commissions with a significantly wider mandate than in the Austrian parliament.

MW: Aus nationalstaatlichen Perspektive könnten so auch Möglichkeiten geschaffen werden, KI rasch genug in die verschiedenen wirtschaftlichen Aktivitäten einzubetten. Das ist ein wichtiges Standbein der FTI-Politik, das weiterbestehen wird. Wie gesagt: Die gesellschaftspolitische Dimension der FTI-Politik wird in den kommenden Jahren eher stärker werden.

Die gesellschaftspolitischen Diskussionen, die dafür notwendig wären, werden derzeit allerdings nicht geführt. Wo sollten diese stattfinden, und wie könnte man sie ankurbeln?

MW: Diese Diskussionen könnten im Umfeld der Parlamente stattfinden, die ja ein zentrales Forum für gesellschaftliche Debatten sein sollten, in denen stellvertretend für die Bürgerinnen und Bürger unterschiedliche politische Zukunftsvorstellungen zur Sprache kommen. Der Deutsche Bundestag beispielsweise führt seit vielen Jahren Debatten zu kontroversen technologischen Themen; es gibt Enquete-Kommissionen, die mit einem deutlich umfassenderen Mandat ausgestattet sind als im österreichischen Parlament.

PSR: Die Diskussion sollte auf breit nutzbare Infrastrukturen fokussieren. Denn Infrastrukturen können strategisch aufgebaut werden. Kartellrecht, Wettbewerbsrecht und Datenschutz sind natürlich wichtig, stellen aber Mechanismen zur Begrenzung der sich gegenwärtig entfaltenden Dynamiken dar. Sie sind reaktiv, nicht proaktiv. Wir brauchen eine proaktive Diskussion, die darauf abzielt, Infrastrukturen aufzubauen, welche die KI der Zukunft tatsächlich praktisch so materialisieren, dass sie allen in der Gesellschaft zugutekommt und keine weitere Spaltung bewirkt.

MW: Eines der Probleme dabei ist, dass dort, wo Regulative entworfen werden, erst Wissen und Kompetenz über KI aufgebaut werden müssen. Wir brauchen viel mehr integrative Expertise, mit der z. B. juristische Fragen so in Angriff genommen werden, dass auch Aspekte der KI aufgegriffen und verarbeitet werden können. Da stoßen wir derzeit massiv an Grenzen: Wir bauen rechtlich auf Prinzipien auf, die in Bezug auf Fragen der KI möglicherweise irrelevant werden.

PSR: The discussion should focus on broadly usable infrastructures. You know, infrastructures can be built up strategically. Antitrust and competition law as well as privacy and data protection are of course important but also are mechanisms to restrict the dynamics that are currently unfolding. They are reactive, not proactive. What we need is a proactive discussion that aims to build infrastructures that make the AI of the future actually and practically materialize in a way that benefits all in society and does not contribute to further division.

MW: One of the problems is that drawing up regulatory provisions first takes building knowledge and competence about AI. We need a lot more integrative expertise that, for example, tackles legal matters in such way as to also take on and process aspects of AI. This is where we are currently facing massive limits: legally, we are building on principles that may become utterly irrelevant with regard to AI issues.

This reminds me of the debate this last spring about the driver service Uber, in which traditional taxi businesses invoked regulations dating from the nineteenth century ...

MW: In the long term, responding to new developments with an outdated set of rules cannot be a solution. Compared to the development of technology in the private sector, the development of institutional framings is way too slow.

AI is a radical and disruptive innovation which, in a Schumpeterian sense, destroys old structures. Can this process be steered in the interest of society?

MW: The question is much less whether the destruction of old structures can, and should, be prevented; old structures will inevitably disappear sooner or later. Rather, it is about whether old structures can be replaced with new ones whose potential can be perceived and developed in a way that is desirable for society. Presently, this whole area eludes the wish for regulation—because it largely is under the control of private businesses from outside Europe. We are not entirely helpless in the face of this disruption, however. There are possibilities of exerting influence such as the principles according to which digital platforms are designed. Isn't it possible to create quicker upscaling mechanisms in Europe as well?

Das erinnert mich an die heuer im Frühjahr geführte Debatte über den Fahrdienst Uber, bei der sich das Taxigewerbe auf Bestimmungen berufen hat, die aus dem 19. Jahrhundert stammen...

MW: Auf Dauer kann es keine Lösung sein, mit einem veralteten Regelwerk auf neue Entwicklungen zu reagieren. Im Vergleich zur technologischen Entwicklung im privatwirtschaftlichen Bereich geht die Entwicklung institutioneller Framings viel zu langsam vor sich.

κ1 ist eine radikale und disruptive Innovation, die im Schumpeter'schen Sinn alte Strukturen zerstört. Kann man diesen Prozess im Dienst der Gesellschaft steuern?

MW: Die Frage ist weniger die, ob man die Zerstörung alter Strukturen verhindern kann oder will; alte Strukturen werden früher oder später unvermeidlich verschwinden. Es geht vielmehr darum, ob man alte Strukturen durch neue ersetzen kann, deren Möglichkeiten in einer Form entwickelt und wahrgenommen werden, die für die Gesellschaft wünschenswert ist. Derzeit entzieht sich dieser Bereich dem Wunsch nach Steuerung – weil er eben einer weitgehend privatwirtschaftlichen Kontrolle von außerhalb Europas unterliegt. Wir sind dieser Disruption aber nicht völlig hilflos ausgeliefert. Es gibt Möglichkeiten der Beeinflussung wie z. B. die Leitprinzipien, nach denen digitale Plattformen gestaltet werden. Können nicht auch in Europa schnellere Upscaling-Mechanismen geschaffen werden?

Eine zentrale Frage der europäischen Politik scheint zu sein, ob die Macht der Oligopole und Monopole gebrochen werden soll oder kann.

PSR: Europa hat bereits einmal Monopolstrukturen auflösen können: Die großen nationalstaatlichen Telekom-Systeme konnten dazu gebracht werden, für sich einen Mehrwert darin zu sehen, dass eine europäische digitale Infrastruktur entsteht. Daran hat die Europäische Kommission in den 1960er-Jahren zu arbeiten begonnen, und das hat letztlich die aktuelle Digitalisierung in Europa ermöglicht. Alle Disruptionen in diesem Bereich – man denke etwa an die Entwicklung von HTML und des World Wide Web am CERN – waren nur durch die Europäisierung der Telekommunikation möglich. Heute sollte es möglich sein, in Europa die Voraussetzung für eine *europäische* κ1 zu schaffen. ✕

One central issue of European politics seems to be whether or not the power of monopolies and oligopolies can, and should, be broken.

PSR: Europe already once succeeded to break up monopolistic structures: the big national telecom companies could be brought to see a surplus value for themselves in having a European digital infrastructure emerge. That is something that the European Commission started working on in the 1960s, and eventually it made the present-day digitization in Europe possible. All disruptions in this area—just think of the development of HTML and the World Wide Web at CERN—were only possible because of the Europeanization of telecommunication. Today, it should be possible to provide in Europe the conditions for a *European AI*. ✕

Was ist Intelligenz, was künstliche Intelligenz?

Die kognitiven Fähigkeiten des Menschen lassen sich kaum in ein Modell pressen. Entsprechend zahlreich sind die Ansätze für eine Definition des Begriffs »Intelligenz«. Noch um einiges schwieriger ist es, den Begriff »künstliche Intelligenz« klar zu fassen. Ein kurzer Streifzug durch die Debatten der vergangenen Jahre und Jahrzehnte.

Das Wort **INTELLIGENZ** leitet sich vom lateinischen »intelligere« (verstehen) ab. Dieses Wort setzt sich aus »inter« (zwischen) und »legere« (lesen, wählen) zusammen, bedeutet also wörtlich »wählen zwischen...«. Wenn von Intelligenz gesprochen wird, ist allerdings selten klar, was genau damit gemeint ist. Es gibt unzählige Definitionen, die sich entweder auf unterschiedliche Aspekte beziehen – etwa auf Lernen oder Problemlösen –, das Phänomen von verschiedenen Blickwinkeln aus betrachten – etwa aus der Perspektive der Entwicklungspsychologie, der Neurobiologie, der Psychiatrie, der Logik oder der Technik – oder unterschiedlich umfassend sind.

Das Onlinelexikon *Wikipedia* schlägt folgende sehr allgemein gehaltene Definition vor: »Intelligenz ist in der Psychologie ein Sammelbegriff für die kognitive Leistungsfähigkeit des Menschen. Da einzelne kognitive Fähigkeiten unterschiedlich stark ausgeprägt sein können und keine Einigkeit besteht, wie diese zu bestimmen und zu unterscheiden sind, gibt es keine allgemeingültige Definition der Intelligenz. Vielmehr schlagen die verschiedenen Intelligenztheorien unterschiedliche Operationalisierungen des alltagssprachlichen Begriffs vor.«¹

Eines der bekanntesten Intelligenzmodelle geht auf den britischen Psychologen Charles Spearman zurück und basiert auf dem sogenannten Generalfaktor der Intelligenz. Dieser »g-Faktor« wird als eine Art Persönlichkeitsmerkmal aufgefasst, der alle Leistungsbereiche beeinflusst. Dieser Auffassung zufolge gibt der Faktor Aufschluss über die geistige Kapazität, die intellektuelle Leistung oder die Verarbeitungsgeschwindigkeit und drückt damit aus, ob eine bestimmte Person eher ein »simpler Charakter« oder ein »begabteres Genie« ist.

Der »g-Faktor« sagt zum Beispiel den schulischen Erfolg, das Prestige des später erreichten Berufs und das Einkommen besser vorher als jede andere erfassbare Eigenschaft. Eine in der Wissenschaft viel diskutierte Frage ist, inwieweit der »g-Faktor« von biologischen (also genetischen) und von sozialen (und wirtschaftlichen) Faktoren geprägt wird – im angloamerikanischen Raum läuft diese Debatte unter dem Schlagwort »nature versus nurture«. Die Vorhersagekraft des Faktors wird deutlich verringert, wenn man den sozioökonomischen Status des Elternhauses mit einbezieht. Ergänzt wird der »g-Faktor«

INTELLIGENZ Alltagssprachlich ist mit Intelligenz die kognitive Leistungsfähigkeit eines Menschen (oder anderer Lebewesen) gemeint. In der Wissenschaft gibt es allerdings keine allgemein anerkannte Definition von Intelligenz – und zwar deshalb, weil es viele theoretische Konzepte gibt, die zum Teil sehr unterschiedliche Ausgangspunkte haben. ✕

¹ <https://de.wikipedia.org/wiki/Intelligenz> (15.05.2018).

What Is Intelligence, and What Is Artificial Intelligence?

Man’s cognitive capacities can hardly be forced into a model. This explains the large number of different approaches to define the concept of “intelligence.” It is even more difficult to clearly grasp the meaning of the term “artificial intelligence.” What follows is a short excursion into the debates of recent years and decades.

The word **INTELLIGENCE** derives from the Latin “intellegere” (understand), which, composed of “inter” (between) and “legere” (read, choose), literally means “choose between...” When the word “intelligence” comes up in a conversation, it is hardly ever clear what exactly is meant by it. There are innumerable definitions of the term that either relate to different aspects (such as learning or solving problems), consider the phenomenon from different perspectives (such as from the angle of developmental psychology, neurobiology, psychiatry, logic, or technology), or vary in their scope. The respective entry in the online encyclopedia *Wikipedia* strikes a rather general note: “Intelligence has been defined in many different ways to include the capacity for logic, understanding, self-awareness, learning, emotional knowledge, reasoning, planning, creativity, and problem solving. It can be more generally described as the ability to perceive or infer information, and to retain it as knowledge to be applied towards adaptive behaviors within an environment or context.”¹

¹ <https://en.wikipedia.org/wiki/Intelligence> (May 15, 2018).

One of the most popular models of intelligence developed by the British psychologist Charles Spearman is based on the so-called “g factor” of general intelligence. This factor is seen as a personality trait that influences all kinds of performance. According to Spearman’s approach, the “g factor” provides a yardstick for a person’s mental capacity, intellectual performance, and processing speed and indicates whether a certain person is a “simple character” or a “more talented genius.” It allows predicting a person’s success at school, the prestige of

INTELLIGENCE In everyday language, intelligence refers to the cognitive capacities of humans (or other living beings). In science, however, there is no generally accepted definition of intelligence—the reason being that there are many theoretical concepts of what makes intelligence, which in part take widely different approaches to the subject. ✘

their later profession, and their income better than any other graspable trait. A matter much under discussion is to what degree the “g factor” is determined by biological (i. e., genetic) and social (and economic) parameters; the catchphrase for this debate in the Anglo-American world is “nature versus nurture.” The predictive power of the factor clearly diminishes when the socioeconomic status of the parental home is taken into account. Spearman also postulated specific secondary intelligence factors he called “s factors” characterizing a person’s mathematical, verbal, and visual skills.

their later profession, and their income better than any other graspable trait. A matter much under discussion is to what degree the “g factor” is determined by biological (i. e., genetic) and social (and economic) parameters; the catchphrase for this debate in the Anglo-American world is “nature versus nurture.” The predictive power of the factor clearly diminishes when the socioeconomic status of the parental home is taken into account. Spearman also postulated specific secondary intelligence factors he called “s factors” characterizing a person’s mathematical, verbal, and visual skills.

bei Spearman durch untergeordnete spezifische Intelligenzfaktoren («s-Faktoren») wie etwa die Leistung bei mathematischen Aufgaben und sprachlichen Problemstellungen oder das räumliche Vorstellungsvermögen.

Intelligenzfaktoren und multiple Intelligenzen

Andere Forscher indes lehnen die Idee eines generellen, übergeordneten Intelligenzfaktors ab und betonen demgegenüber die bereichsspezifische Organisation der Intelligenz. Demnach wird Intelligenz als die Summe der Leistungsfähigkeit in verschiedensten Teilbereichen der Kognition aufgefasst. Ein häufig verwendetes Modell geht auf den us-Ingenieur und Psychologen Louis Leon Thurstone zurück. Dieser unterscheidet sieben Primärfaktoren der Intelligenz: die Leistung bei räumlich-visuellen Aufgaben, die Wahrnehmung von Objekten und Beziehungen zwischen diesen, rechnerisch-mathematische Fähigkeiten, die Gedächtnisleistung, logisches Schlussfolgern, Wortflüssigkeit und das richtige Interpretieren verbaler Beziehungen.

Manche Forscher, unter ihnen der us-Erziehungswissenschaftler Howard Gardner, gehen noch einen Schritt weiter und postulieren sogar voneinander unabhängige Intelligenzen («multiple Intelligenzen»), die auch im Gehirn als getrennte Organisationseinheiten verortet sind. Fällt eine aus, so hat das dieser Theorie zufolge keine Auswirkung auf die anderen Intelligenzen. Wieder andere Forscher, etwa der britische Persönlichkeitspsychologe Raymond Bernard Cattell, haben Modelle mit nur zwei Intelligenzfaktoren konstruiert: auf der einen Seite eine »fluide Intelligenz«, die vererbt wird und die beispielsweise die geistige Kapazität, die Auffassungsgabe oder das generelle Verarbeitungsniveau bestimmt; auf der anderen Seite eine »kristalline Intelligenz«, die alle Fähigkeiten umfasst, die im Lauf des Lebens erlernt und durch die Umwelt bestimmt werden – Intelligenz bezieht sich hier vor allem auf die Fähigkeit, dieses erworbene Wissen anzuwenden.

Wie man Intelligenz misst

Die kognitive Leistungsfähigkeit wird mithilfe von Intelligenztests gemessen. Das Ergebnis wird häufig in Form eines **INTELLIGENZQUOTIENTEN** (IQ) angegeben: Dabei wird angenommen, dass die Intelligenz in der Bevölkerung normalverteilt ist; als Mittelwert wird ein IQ von 100 definiert (mit einer Standardabweichung von 15). Ein konkretes Testergebnis gibt somit die Abweichung von diesem Mittelwert (für eine bestimmte Altersklasse) an. Nach der »klassischen« Definition des ersten IQ-Tests der französischen Psychologen Alfred Binet und Théodore Simon im Jahr 1904 ergibt sich der IQ aus dem mit 100 multiplizierten Quotienten von Intelligenzalter und Lebensalter.

Da es viele unterschiedliche Sichtweisen auf Intelligenz gibt, gibt es auch ebenso viele unterschiedliche Testverfahren. Diese wurden für bestimmte Anwendungen konzipiert – etwa zur Feststellung der Schulreife, für forensische Gutachten, zur Quantifizierung geistiger Behinderungen oder zur Feststellung der Eignung für bestimmte Berufe. Angesichts dieser Vielfalt meint ein geflügeltes Wort (das auf

INTELLIGENZQUOTIENT Viele Konzepte von Intelligenz beinhalten eine Messmethode, mit der die kognitive Leistungsfähigkeit eines Individuums getestet werden kann. Das Ergebnis wird häufig in Form eines Intelligenzquotienten (im Verhältnis zum Durchschnitt) ausgedrückt. Intelligenz sei genau das, was ein IQ-Test misst, lautet ein geflügeltes Wort in der Psychologie. ✕

Intelligence factors and multiple intelligences

Other scientists reject the idea of a general, overriding intelligence factor and emphasize the area-specific organization of intelligence. They define intelligence as the sum of abilities in various specific fields of cognition. The US engineer and psychologist Louis Leon Thurstone conceived a frequently used model. It distinguishes between seven primary mental abilities: spatial visualization, perceptual speed, numerical ability, rote memory, reasoning, word fluency, and verbal comprehension.

Some scholars such as the US pedagogue Howard Gardner go one step further and postulate intelligences independent of each other (“multiple intelligences”), which are also located in the brain as separate organizational units. According to this theory, the breakdown of one of these intelligences does not affect the others. Other scholars such as the British personality psychologist Raymond Bernard Cattell developed a model based on only two intelligence factors: “fluid intelligence,” which is hereditary and determines a person’s mental capacity, perceptive faculty, or general level of processing, and “crystallized intelligence,” which encompasses all skills acquired in the course of an individual’s life and determined by their environment; this form of intelligence is mainly concerned with applying the knowledge acquired.

Measuring intelligence

A person’s cognitive capacity is measured by means of intelligence tests. The result is frequently given in the form of an **INTELLIGENCE QUOTIENT** (IQ). Intelligence is assumed to be equally distributed among the population; the median raw score of the norming sample is defined as IQ 100 (with a standard deviation of 15 IQ points greater or less). Thus, each test result renders the deviation from this median value (for a certain age group). Following the historical definition of the first IQ test carried out by the French psychologists Alfred Binet and Théodore Simon in 1904, the intelligence quotient is obtained by dividing a person’s mental age score by her or his chronological age and multiplying the fraction by 100.

Since there are many different views of what intelligence is, we also find many different intelligence tests. The tests were conceived for specific

applications such as assessing a child’s school readiness, forensic expertise, quantifying mental handicaps, or evaluating a person’s qualification for a special job. This great variety may be behind the saying (coined by the experimental psycholo-

gist Edwin Boring) that intelligence is what an IQ test measures—a statement that reveals a rather pragmatic approach to the issue of defining intelligence.

INTELLIGENCE QUOTIENT Many concepts of intelligence include a measurement method to test and assess an individual’s cognitive capacity. The result is frequently expressed in the form of a so-called intelligence quotient (as compared to the human average). A common tenet in psychology is that intelligence is what is measured by an IQ test. ✖

den us- Experimentalpsychologen Edwin Boring zurückgeht), dass Intelligenz genau das sei, was ein IQ-Test misst – was einen ziemlich pragmatischen Zugang zum Problem der Definition von Intelligenz darstellt.

Emotionale und soziale Intelligenz

Neben den bisher beschriebenen Ansätzen, mit denen die kognitive Leistungsfähigkeit erklärt und gemessen werden soll, tauchten in jüngerer Zeit weitere Intelligenzbegriffe auf, die für die Fähigkeiten des Menschen ebenfalls wichtig sind. So wurde im Jahr 1990 von den us-Psychologen John D. Mayer und Peter Salovey der Terminus »emotionale Intelligenz« eingeführt. Dieser beschreibt die Fähigkeit des Menschen, eigene und fremde Gefühle korrekt wahrzunehmen, zu verstehen und zu beeinflussen – was ebenfalls ein wichtiger Faktor ist, um die oben erwähnte Frage nach dem Erfolg im Leben und im Beruf beantworten zu können.

Ebenso wichtig dafür ist der in jüngster Zeit in Mode gekommene Begriff der »sozialen Intelligenz«. Darunter versteht man einen »Komplex von Fähigkeiten, die dazu dienen, in Kommunikations- und Interaktionssituationen entsprechend den Bedürfnissen der Beteiligten Realitätskontrolle zu übernehmen und effektiv zu handeln«, wie es in Dorschs *Lexikon der Psychologie* heißt.² Wie grundlegend dieser Faktor ist, ist schon seit den 1920er-Jahren klar, als der us-Psychologe Edward Lee Thorndike bemerkte, dass der (fachlich) beste Mechaniker als Vorarbeiter scheitern werde, wenn es ihm an sozialer Intelligenz fehle.

² Markus Antonius Wirtz (Hg.), *Dorsch – Lexikon der Psychologie*, 18., überarb. Aufl., Bern 2017; <https://portal.hogrefe.com/dorsch/de/startseite/> (15.05.2018).

Was ist künstliche Intelligenz?

Mindestens genauso schwammig wie die Definition der menschlichen Intelligenz ist die von **KÜNSTLICHER INTELLIGENZ** (KI). Es gibt auch hier keinen allgemein akzeptierten Begriff. Das spiegelt sich auch in diesem Jahrbuch wider; in den verschiedenen Beiträgen finden sich ganz unterschiedliche Auffassungen. Manche meinen mit KI die Fähigkeit, Probleme zu lösen oder komplexe Ziele zu erreichen. Andere betrachten ein System dann als intelligent, wenn es Fähigkeiten hat, die eigentlich nur dem Menschen zukommen. Wieder andere Experten sehen das logische Denken oder das Lernen als zentral an. Manche rücken Aspekte wie Verstehen oder Entscheidungsfindung in den Mittelpunkt.

Die *Wikipedia* definiert KI als »Teilgebiet der Informatik, welches sich mit der Automatisierung intelligenten Verhaltens und dem Maschinenlernen befasst«.³ Diese Definition ist natürlich insofern schwach, als unklar ist, was mit »intelligentem Verhalten« genau gemeint ist. Auch einer der Väter des Begriffs KI, der us-Forscher Marvin Minsky (der den Terminus 1956 bei der Dartmouth-Konferenz mitgeprägt hat), war sich nicht ganz im Klaren, was genau damit bezeichnet wird. Heute plädieren viele Forscher dafür, den Begriff KI gar nicht mehr zu verwenden, sondern stattdessen pragmatisch von Maschinenlernen zu sprechen.

KÜNSTLICHE INTELLIGENZ Künstliche Intelligenz ist laut *Wikipedia* ein »Teilgebiet der Informatik, welches sich mit der Automatisierung intelligenten Verhaltens und dem Maschinenlernen befasst«. Dabei bleibt offen, was mit »intelligentem Verhalten« gemeint ist (für das es keine allgemeingültige Definition gibt). Viele Forscher meinen daher, dass man den Begriff KI nicht mehr verwenden, sondern besser von »Maschinenlernen« sprechen sollte. ✕

³ https://de.wikipedia.org/wiki/Künstliche_Intelligenz (15.05.2018).

Emotional and social intelligence

Next to the approaches described so far, which are supposed to explain and measure cognitive capacity, recently presented concepts of intelligence are aimed at not less important human qualities. In 1990, for example, US psychologists John D. Mayer and Peter Salovey introduced the term “emotional intelligence.” This term describes the capability of individuals to correctly recognize their own emotions and those of others, understand, and influence them—which is also an important factor when dealing with a person’s success in her or his life and job.

The concept of “social intelligence,” which has lately come into fashion, is equally important for such assessments. This kind of intelligence is seen as a complex of capacities that help to control reality in communicating and interacting with others and act effectively according to the involved people’s demands, as Dorsch’s *Lexikon der Psychologie* enlightens its readers.² There has been no doubt about the significance of this factor since the 1920s, after the US psychologist Edward Lee Thorndike had observed that the best (trained) mechanic would fail as a foreman when he lacked social intelligence.

² See Markus Antonius Wirtz, ed., *Dorsch – Lexikon der Psychologie*, 18th, revised ed., Bern 2017; <https://portal.hogrefe.com/dorsch/de/startseite/> (May 15, 2018).

What is artificial intelligence?

The definition of **ARTIFICIAL INTELLIGENCE** (AI) is as woolly as that of human intelligence. There is no generally accepted concept in this case either. This fact is also reflected in this yearbook; the various contributions reveal quite different attitudes. Some of the authors understand AI as the skill to solve problems and achieve complex goals. Others regard a system as intelligent if it shows capacities that are generally associated with the nature of man. And yet other experts consider logical thinking or learning the crucial factors. Some focus on understanding or decision-finding.

Wikipedia defines AI research as “the study of ‘intelligent agents’: any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals.”³ This is a poor definition, as it does not clarify what “intelligent” means. One of the fathers of the term AI, the US scholar Marvin Minsky (who was one of those who coined the expression at the Dartmouth Workshop in 1956), was also not sure about what AI exactly means. Today, many scholars argue that it would be best to do without the term at all and pragmatically rely on the expression “machine learning” instead.

³ https://en.wikipedia.org/wiki/Artificial_intelligence (May 15, 2018).

ARTIFICIAL INTELLIGENCE According to *Wikipedia*, artificial intelligence is “intelligence demonstrated by machines,” which is “when a machine mimics ‘cognitive’ functions...such as ‘learning’ and ‘problem solving.’” Given that there is no generally accepted definition of what intelligence actually is, many researchers have suggested abandoning the term AI and speaking of “machine learning” instead. ✖

Starke und schwache KI

Etwas mehr Klarheit in die Sache bringt eine begriffliche Unterscheidung, die der us-Philosoph John R. Searle vor rund 40 Jahren eingeführt hat. Er sprach auf der einen Seite von »starker KI«: Damit meint er die technische Nachbildung der allgemeinen menschlichen Intelligenz, die ein universelles Werkzeug zum Lösen von Problemen in praktisch allen Teilbereichen darstellt. Hierher gehören viele der heißumstrittenen Fragen wie etwa, ob Computer Menschen irgendwann überflügeln werden (»Singularität«), ob Maschinen so etwas wie ein Bewusstsein entwickeln können oder ob Cyborgs eine nächste Stufe der Evolution sein könnten.

Der »starken KI« stellte Searle die »schwache KI« gegenüber: In diesem Fall geht es um die Entwicklung konkreter technischer Systeme, die bestimmte wohldefinierte Aufgaben übernehmen können, die bisher dem Menschen vorbehalten waren. Die »schwache KI« zielt also »nur« darauf ab, spezialisierte Maschinen zu bauen, die den Menschen unterstützen und ihm Arbeit abnehmen. Dazu zählen etwa das Suchen von Informationen in riesigen Datenbanken, das Klassifizieren von Objekten, das Erkennen von Mustern und Korrelationen, Spracherkennung, Übersetzungsprogramme, die intelligente Steuerung von Industrierobotern (Industrie 4.0), die Entwicklung von autonomen Autos oder der Aufbau von Expertensystemen. Der Schwerpunkt der »schwachen KI« liegt auf Anwendungen, die es schon gibt oder die in absehbarer Zeit funktionieren könnten. Die philosophischen Probleme einer derzeit nicht realisierbaren »allgemeinen KI« spielen dabei kaum eine Rolle.

Unterschiede zwischen Mensch und Maschine

Hilfreich beim Verständnis von KI ist, wenn man sich einige grundlegende Unterschiede zwischen menschlicher Intelligenz und künstlichen intelligenten Systemen vor Augen führt (und dabei stets im Hinterkopf behält, dass einige der Unterschiede zunehmend geringer werden, doch in absehbarer Zeit nicht verschwinden werden).

Das Denken des Menschen funktioniert auf biologischer Basis (also auf der Grundlage von Zellen, DNA, Botenstoffen, Proteinen usw.), Computer hingegen können Informationen dank der besonderen Eigenschaften von Halbleitermaterialien verarbeiten. Das Gehirn besteht aus rund 80 bis 100 Milliarden Nervenzellen; jedes Neuron ist über Synapsen im Schnitt mit 1000 anderen Nervenzellen verknüpft. Die Strukturen im Gehirn verändern sich ständig (darin liegt wahrscheinlich das Wesen des Lernens), wohingegen Computer fix verdrahtet sind. Überdies arbeitet die Biologie ungemein effizient: Alle kognitiven Tätigkeiten zusammen erfordern eine Leistung von gerade einmal 50 Watt – wohingegen ein Supercomputer mit einer vergleichbaren Anzahl von Transistoren mehrere Räume füllt und so viel Energie verbraucht wie eine ganze Siedlung. Dafür sind Computer wesentlich schneller, und die Leistungsfähigkeit steigt rasant weiter: Dem **MOORESCHEN GESETZ** zufolge verdoppelt sich diese alle eineinhalb bis zwei Jahre.

MOORESCHES GESETZ 1965 postulierte der spätere Intel-Gründer Gordon E. Moore, dass sich die Anzahl an Transistoren, die in einen integrierten Schaltkreis festgelegter Größe passen, jedes Jahr verdoppelt. Später veränderte er den Verdopplungszeitraum auf alle zwei Jahre, heute gelten eineinhalb Jahre als jene Größe, welche die exponentielle Entwicklung der Rechenleistung in den vergangenen 50 Jahren richtig beschreibt. Ob das »Gesetz« auch für die Zukunft gilt, ist derzeit offen ✕

Strong and weak AI

A conceptual distinction introduced by the us philosopher John R. Searle about forty years ago sheds a little more light on the matter. He distinguished between “strong AI” and “weak AI.” He understood the former as the technical reconstruction of general human intelligence that provided a universal instrument for solving problems in virtually all spheres. This concerns many hotly debated issues such as whether computers will surpass humans one day (“singularity”), whether machines will be able to develop a kind of consciousness, or whether cyborgs will be an element of the next step of evolution.

“Weak AI,” on the other hand, is about the development of specific technological systems that can take on well-defined tasks reserved for man to date. “Weak AI” is “merely” aimed at building specialized machines that support man, relieving humans of certain chores. These include the search for information in huge databases, the classification of objects, the recognition of patterns and correlations, speech recognition, translation programs, the intelligent control of industrial robots (Industry 4.0), the development of autonomous cars, and the establishment of expert systems. The focus of “weak AI” is on applications that already exist and might work in the near future. The philosophical problems of a “general AI” that cannot be realized at the moment are not relevant in this context.

Differences between man and machine

If we want to understand what AI is, it helps to keep some fundamental differences between human intelligence and artificial intelligence systems in mind (and to remember that some differences are on the wane, yet will not vanish in the near future).

Whereas human thinking is based on biological givens (i. e., on cells, DNA, semiochemicals, proteins, etc.), computers process information thanks to the properties of semiconductor materials. The human brain

comprises about 80 to 100 billion neurons; on an average, each neuron is linked to 1,000 other neurons through synapses. Structures in the human brain undergo a continuous change (which probably reflects the nature of learning), whereas computers are permanently wired.

MOORE'S LAW In 1965, Gordon E. Moore, who later cofounded Intel, contended that the number of transistors on a dense integrated circuit of a defined size doubles every year. Later, he extended that period to two years; today, one and a half years are considered to be the measure that correctly describes the exponential development of computing performance over the past fifty years. Whether Moore's “law” will also apply in the future still remains to be seen. ✕

Biology works in an enormously efficient way: All cognitive activities together do not require more than 50 watts of power, while a supercomputer with a comparable number of transistors fills several rooms and spends as much energy as an entire residential estate. Computers, however, are much faster, and their capacity increases rapidly; according to **MOORE'S LAW** it doubles every one and a half to two years.

Die dichte Vernetzung im Gehirn ermöglicht es, dass die Verarbeitung von Sinnesreizen und das Entstehen von Erregungsmustern hochgradig parallel ablaufen. Im Gehirn gibt es keine die Prozesse steuernde Instanz. Das Gehirn kennt auch keine speziellen Orte, an denen Informationen gespeichert werden. Ganz anders bei Computern, zumindest bei jenen mit heutiger Architektur: Diese sind hierarchisch aufgebaut und verfügen über separate Bauteile für die Speicherung, Koordinierung und Verarbeitung der Daten – die im Wesentlichen seriell abgearbeitet werden. Durch den Bau von Rechnern mit mehreren Zentraleinheiten (cpus) und den vermehrten Einsatz von Grafikkarten beginnt sich das freilich zu ändern; im Vergleich zum Gehirn ist die Parallelisierung in der Informatik allerdings immer noch sehr gering. Einen klaren Fortschritt in der Computertechnik bringen die aktuellen neuronalen Netzwerke mit mehreren Ebenen, die es auf Milliarden von Verknüpfungen zwischen den Millionen simulierten Neuronen bringen können. Dennoch bleibt der Vernetzungsgrad im Vergleich zum Gehirn weiterhin ziemlich bescheiden.

Wie Informationen verarbeitet werden

Die Art, wie Informationen verarbeitet werden, ist in Computern grundlegend anders als im Gehirn. Computer punkten damit, dass sie immens schnell sind. Die Natur hat indes viele Verfahren entwickelt und perfektioniert, mit denen Informationen auf sinnvolle Weise verdichtet werden. So liefern beispielsweise Sinnesorgane nicht alle Daten an das Gehirn weiter, sondern filtern schon im Vorfeld die relevanten Informationen heraus.

Computer funktionieren klassischerweise mit Wenn-Dann-Regeln. Programmierer entwerfen nach der exakten Analyse eines Problems einen Algorithmus, der dieses lösen kann. Das funktioniert freilich nur bei Themenstellungen, die sich vollständig in Einzelschritte zerlegen lassen und bei denen es eine eindeutige Zuordnung zwischen Input und Output gibt. Computer sind mittlerweile auch schon sehr gut, wenn es gilt, aus großen Datenmengen statistische Schlüsse – etwa Korrelationen – zu bilden. Auch das beruht auf exakt definierten Rechenvorschriften.

Das Gehirn folgt hingegen anderen Regeln. Auch wenn die Hirnforschung noch bei Weitem nicht so weit ist, alle Abläufe im Gehirn vollständig erklären zu können, so sind dennoch bereits einige Grundprinzipien klar. So postuliert beispielsweise das »Liquid Computing Model«, dass bestimmte Schaltkreise nicht nur für eine einzige Aufgabe zuständig sind, sondern mehrere Aufgaben übernehmen können. Im Gehirn gibt es zudem offenbar Mechanismen, durch die ständig Modelle und mögliche Bilder der Wirklichkeit produziert werden, von denen die erfolgreichen – also jene, die zumindest Teile der Welt korrekt beschreiben – weiter benutzt werden. Diese Vorgänge passieren parallel, blitzschnell und völlig unbewusst und machen es möglich, dass wir uns scheinbar mühelos in der Welt bewegen können.

Owing to the dense interconnectedness within the human brain, the processing of sensory stimuli and the genesis of agitation patterns take place almost simultaneously. The brain knows no authority controlling the various processes, neither are there any special places where information is stored. Computers, at least those with today's architecture, are quite different in this regard. They are hierarchically structured and comprise separate components for storage, coordination, and processing of data—tasks that are mainly serially dealt with. This has begun to change through the construction of computers with several central processing units (CPUs) and the increased use of graphics cards; compared to the human brain, however, parallel computing is still a very rare approach. Current neural networks with several levels, which may be made up of millions of simulated neurons with billions of connections between them, have brought about a marked advance in computer technology. Compared to the human brain, such networks' degree of interconnectedness is still rather modest.

How information is processed

Computers and the human brain process information fundamentally differently. Computers score with their enormous speed. Yet nature has developed and perfected numerous procedures in which information is condensed in a constructive way. Sense organs do not forward all data to the brain, for example; they already filter the relevant information out beforehand.

Computers traditionally rely on if-then rules. After precisely analyzing a problem, programmers design an algorithm that is able to solve this problem. This only works with issues that can be split up into individual steps and allow an unequivocal correlation between input and output. In the meantime, computers have become very good at forming statistical conclusions such as correlations from huge amounts of data. This progress is also based on exactly defined algorithms.

The human brain, however, follows different rules. Even if neuroscience is clearly far from being able to fully explain the processes within the human brain, some principles have already become clear. The Liquid Computing Model, for instance, postulates that certain circuits are not responsible for only a single task but may take on several tasks. The brain comprises mechanisms that continuously produce models and possible images of reality; the ones that are successful, i. e., describe at least parts of the world correctly, will be accepted for further use. These processes happen in parallel, with lightning speed and completely unconsciously, and enable us to move in the world in a seemingly effortless way.

Deterministisch oder kreativ?

Die Kreativität, zu der Menschen fähig sind, beweist, dass die Vorgänge im Gehirn nicht völlig deterministisch ablaufen. Herkömmliche Computersysteme inklusive der »klassischen« Methoden der KI (Algorithmen, Statistiken) sind hingegen völlig deterministisch. Dieser Unterschied zwischen Mensch und Maschine beginnt sich nun aufzuweichen: In komplexen künstlichen neuronalen Netzwerken, auf denen Deep Learning beruht, werden nichtdeterministische Vorgänge realisierbar, die keine eindeutigen Wenn-Dann-Beziehungen mehr sind. Was diese Verfahren allerdings zurzeit noch nicht können – was unser Gehirn aber mühelos schafft – ist, neue Erkenntnisse in einen Kontext einzufügen.

Bei künstlichen neuronalen Netzen kommt es immer wieder vor, dass sie falsche Unterscheidungsmerkmale lernen, weil sie den Kontext nicht kennen. Ein Beispiel dafür: Nachdem man einem neuronalen Netzwerk zahlreiche Bilder mit Panzern vorgelegt hatte, wurden diese in den Trainingsdaten zu einem sehr hohen Prozentsatz richtig erkannt. Bei anderen Bildern funktionierte das aber plötzlich nicht mehr. Es stellte sich heraus, dass sich der Lernalgorithmus als Klassifikationsmerkmal nicht die Form des Panzers, sondern die Tatsache angeeignet hatte, dass im Hintergrund der Bilder Bäume zu sehen waren (die Fotos der Panzer waren alle im Freien geschossen worden).

Für die Informatik der Zukunft sind Erkenntnisse darüber essenziell, **WIE SICH MENSCHEN NEUES WISSEN ANEIGNEN** – etwa wie Kinder ein Verständnis für die sie umgebende Welt entwickeln. Hier sind zwei Zugänge zu unterscheiden: Bei einem Top-down-Ansatz werden aus bisherigen Erfahrungen bestimmte Merkmale abstrahiert; daraus wird ein Modell der Wirklichkeit konstruiert und bei nächster Gelegenheit getestet, wie gut es die Wirklichkeit beschreibt (das machen beispielsweise Spamfilter in E-Mail-Programmen). Die zweite Möglichkeit ist ein Bottom-up-Ansatz: In diesem Fall werden aus einer Unzahl von Beispielen bestimmte Regelmäßigkeiten oder Muster abgeleitet (wie es Deep-Learning-Systeme machen). Beim Menschen kommen laut aktuellen Erkenntnissen der Entwicklungspsychologie beide Prinzipien zum Einsatz, wobei Kinder beim Lernen offenbar die Stärken beider Verfahren kombinieren.⁴ Die Informatik will die beiden Ansätze in Zukunft in KI-Systemen kombinieren. Das Ergebnis sollen wesentlich leistungsfähigere Systeme sein, die beim Lernen auch Kontextwissen berücksichtigen könnten.

Wie intelligent ist KI?

Aus dieser vergleichenden Betrachtung von menschlicher und künstlicher Intelligenz ist unschwer zu erkennen, dass heutige KI-Systeme zwar bei einer Reihe von Spezialproblemen überlegen sein können, dass der Mensch aber unterm Strich weiterhin haushoch überlegen ist – eben weil er über eine umfassende allgemeine Intelligenz verfügt und KI-Systeme (im Sinn von Searles Unterscheidung) »schwach« sind.

POLANYIS PARADOXON 1966 prägte der ungarisch-britische Forscher Michael Polanyi den Satz: »Wir wissen mehr, als wir sagen können.« Dieses implizite Wissen (tacit knowledge), über das wir Menschen verfügen, lässt sich kaum formalisieren und auf Computer übertragen. Viele Aufgaben, die uns leichtfallen, sind für Maschinen daher nur schwer lösbar. Das beginnt sich nun durch maschinelle Lernverfahren zu ändern. ✕

⁴ Siehe Alison Gopnik, »Lernen wie ein Kind«, in: *Gehirn & Geist*, Jg. 2018, Nr. 5, S. 51–57.

Determinist or creative?

The creativity man is capable of shows that processes in the human brain do not take place completely deterministically. Traditional computer systems including the “classical” AI methods (algorithms, statistics) are entirely determinist, however. This difference between man and machine begins to blur now. Non-determinist processes that no longer represent unequivocal if-then relations can be realized in complex artificial neural networks, which deep learning is based on. Yet these processes are not able to introduce new insights into a context—which is something the human brain does without effort.

Artificial neural networks frequently pick up wrong distinctive features because they are not familiar with the context. The following example illustrates the problem: After numerous pictures of tanks had been presented to a connectionist system, these tanks were correctly identified to a very high degree in the training data. All at once, the network did not work correctly when confronted with different pictures. It turned out that the algorithm had not learned the shape of the tank as a classification characteristic but the fact that there were trees in the background (all pictures had been taken outdoors).

Future computer science will essentially depend on insights into **HOW HUMAN BEINGS ACQUIRE NEW KNOWLEDGE**—how children develop an understanding of their surroundings, for example. Two strategies are to be distinguished in this context: When pursuing a top-down approach, certain features are abstracted from made experiences; these features provide the basis for the construction of a model of reality, which is put to the test at the next opportunity as to how far it describes the reality in question (spam filters of e-mail programs do this, for example). The second strategy is a bottom-up strategy: Certain regularities or patterns are derived from innumerable examples (which is something deep learning systems do). According to the latest insights of developmental psychology, the human mind employs both strategies; children who learn obviously combine the fortes of both approaches.⁴ Computer scientists intend to combine both procedures in AI systems in the future. The plan is to arrive at much more efficient systems that also incorporate contextual knowledge in their learning process.

⁴ See Alison Gopnik, “Lernen wie ein Kind,” *Gehirn & Geist*, vol. 2018, no. 5, 51–57.

How intelligent is AI?

This comparison between human and artificial intelligence clearly indicates that today’s AI systems may be superior in solving a number of special problems, but that humans are still miles ahead—because they are equipped with a comprehensive general intelligence and AI systems are “weak” in the sense of Searle’s distinction.

POLANYI’S PARADOX In 1966, the Hungarian-British scholar and polymath Michael Polanyi formulated the insight that “We know more than we can tell.” This tacit knowledge that we humans have can hardly be formalized and transferred to computers. Many tasks that we find easy therefore are difficult to solve for machines. This is beginning to change due to machine-learning processes. ✖

Dennoch stellt sich die Frage: Wie könnte man erkennen, dass ein Computer einem Menschen intellektuell zumindest ebenbürtig oder sogar überlegen ist? Der britische Pionier der Informatik Alan Turing hat 1950 dafür eine Versuchsanordnung konzipiert: Bei diesem **TURING-TEST** führt ein menschlicher Fragesteller ausschließlich über eine Tastatur und einen Bildschirm mit zwei ihm unbekanntem Gesprächspartnern eine Unterhaltung. Der eine Gesprächspartner ist ein Mensch, der andere eine Maschine. Wenn der Fragesteller nach einer intensiven Befragung nicht klar sagen kann, welcher von beiden die Maschine ist, hat die Maschine den »Turing-Test« bestanden, der nun ein dem Menschen ebenbürtiges Denkvermögen unterstellt werden kann. In einzelnen abgegrenzten Bereichen haben manche Maschinen diesen Test bereits bestanden. Der »Turing-Test« hat heute noch seine praktische Anwendung im Captcha, bei dem verzerrte optische Zeichen bei Login-Vorgängen korrekt erkannt werden müssen. Damit wird bei Internetapplikationen überprüft, ob sich gerade ein Mensch und nicht eine Maschine einloggen will.

TURING-TEST 1950 formulierte Alan Turing eine Idee, wie man feststellen könnte, ob ein Computer ein dem Menschen gleichwertiges Denkvermögen hätte: Man stellt Fragen, und wenn man aus den Antworten nicht erkennen kann, ob sie von einem Computer oder von einem Menschen kommen, wird der Maschine ein dem Menschen ebenbürtiges Denkvermögen unterstellt. Der »Turing-Test« prüft nur Funktionalität, nicht das Vorhandensein von Intentionalität oder Bewusstsein. ✘

Im Allgemeinen ist der Unterschied zwischen menschlicher und künstlicher Intelligenz jedoch weiterhin groß. Chinesische Forscher haben kürzlich sogar versucht, diesen zu beziffern: Sie haben regelrechte Intelligenztests mit öffentlich zugänglichen KI-Systemen wie dem von Google oder Siri von Apple durchgeführt. Die Vorgangsweise mag zwar zweifelhaft sein, weil Äpfel und Birnen miteinander verglichen werden. Doch das Ergebnis ist dennoch interessant: Die KI-Systeme erreichen einen IQ von höchstens 47 – was ungefähr dem eines sechsjährigen Kindes im Vergleich zu dem eines durchschnittlich intelligenten Erwachsenen entspricht. Die Fortschritte sind allerdings gewaltig: Ähnliche Tests wurden auch schon im Jahr 2014 durchgeführt. Damals kam das beste KI-System auf einen IQ von 26.⁵ ✘

⁵ »Intelligenztest für Siri«, in: *Technology Review*, 19.10.2017; <https://www.heise.de/tr/artikel/Intelligenztest-fuer-Siri-3864492.html> (15.05.2018).

Nevertheless, the question arises how we might ascertain whether a computer is intellectually at least on a par with—or even superior to—man. The British pioneer of computer science Alan Turing conceived an experimental set-up for this issue in 1950. The **TURING TEST** has a human evaluator, who relies exclusively on a keyboard and a screen, hold a conversation with two partners he does not know. One of the partners is a human, the other a machine. If, after an intense interview, the evaluator is not able to state clearly which of the two is a machine, the machine will have passed the test and it can be assumed that its intellectual capacity is on a level with that of the human mind. Some machines have already passed the test in specific clearly defined areas. The Turing test still finds its practical application in the CAPTCHA test (Completely Automated Public Turing test to tell Computers and Humans Apart), in which distorted letters or digits must be correctly read as part of a login process. Used in Internet applications, this procedure checks whether a human and not a machine is about to log in.

Generally, the difference between human and artificial intelligence is still substantial. Chinese scientists have recently even tried to put a number on this difference by subjecting generally accessible artificial intelligence systems like Google's AI or Apple's Siri to regular intelligence tests. The approach may be doubtful, as it compares apples to oranges. The result is interesting, however: AI systems scored an IQ of 47 at the most—which corresponds to that of a six-year-old child compared to that of an adult of average intelligence. Yet the steps forward are huge. Similar tests were already carried out in 2014: the best system came off with an IQ of 26.⁵ ✕

⁵ "Intelligenztest für Siri," *Technology Review*, October 19, 2017; <https://www.heise.de/tr/artikel/Intelligenztest-fuer-Siri-3864492.html> (May 15, 2018).

TURING TEST In 1950, Alan Turing suggested an experiment to determine whether a computer can have a capacity for thinking equal to that of a human: you ask questions, and if you cannot tell from the answers whether they come from a computer or a human, the machine can be said to have thinking equivalent to that of a human. However, the Turing test is only about functionality, it says nothing about the presence of intentionality or consciousness. ✕



S
S
A

»
K
se
nt
P
zi
ü
d
m
u
m
ft
a
V
S
»l
e
w
a
E
w
O

«
S
A
fe
h
S
a
m
si
a
m
b
o
ir
a
ci
«f
a
ir
le
a

—
©
tc

Samantha
Sergi Santos, Synthea
Amatus (SL/ES)

»Samantha« ist eine mit KI ausgestattete Robotersexpuppe, die auf Berührungen reagiert. Die Puppe mag es, berührt zu werden, und verfügt über verschiedene Modi der Interaktion wie »romantisch«, »familiär« und »sexy«. Samantha möchte auf ihre voll funktionsfähigen Lippen, auf ihre Brüste und ihre Vagina geküsst werden. So kann sie aus dem »Familienmodus« zu einem Punkt gebracht werden, an dem sie auf einer sexuellen Ebene interagieren will, was sogar bis zum Orgasmus gehen kann.

“Samantha” is a robot sex doll equipped with AI that responds differently according to how she is treated. She likes to be touched and has different modes of interaction, such as a “romantic,” a “family” and a “sexy” mode. She wants to be touched and kissed on her fully functioning lips, her breasts, and her vagina to change her mode from “family” to the point at which she wants to interact on a sexual level until she even has an orgasm.

© Ars Electronica/
tom mesic



Geschichte der künstlichen Intelligenz

Maschinen mit menschenähnlichem Verhalten haben schon in der Barockzeit fasziniert. Künstliche Intelligenz im engeren Sinn ist in den ersten Jahren nach dem Zweiten Weltkrieg entstanden; ihren Namen bekam die Technologie im Jahr 1956. Nach Jahren des grenzenlosen Optimismus rutschte die Forschung in den 1970er-Jahren allerdings in einen 20 Jahre dauernden »KI-Winter« – bevor sie dann umso machtvoller wieder auferstand.

Man schrieb den Frühling des Jahres 1770. Am Wiener Hof herrschte große Aufregung und Begeisterung, nachdem der österreichisch-ungarische Hofbeamte und Mechaniker Wolfgang von Kempelen Kaiserin Maria Theresia sein neuestes Werk vorgeführt hatte: einen Schachtisch, an dem eine menschenähnliche Figur saß, die per mechanischem Arm Schachfiguren bewegen konnte – und gegen die meisten Gegenspieler auch gewann. Da die Figur türkisch gewandet war, wurde der **APPARAT** schnell »Schachtürke« genannt. Man rätselte lange Zeit, wie die Maschine, in deren Innerem deutlich hörbar Zahnräder am Werk waren, dies schaffte. Der Erfinder, der das Geheimnis Zeit seines Lebens bewahrte, erntete auf ausgedehnten Europatourneen Ruhm und Ehre. Viele Beobachter vermuteten einen Betrug: Wie sollte denn auch ein Uhrwerk an die geistigen Leistungen des Menschen herankommen – in einer Zeit, in der auch Tiere als bloße Maschinen angesehen wurden? Andererseits hatten frühe Aufklärer wie etwa der französische Philosoph Julien Offray de La Mettrie bereits die Idee in die Welt gesetzt, dass auch Menschen nichts anderes als Maschinen seien. Also zeigte sich so mancher Zeitgenosse von Kempelens Apparat überzeugt.

Wie dem auch sei: Jahrzehnte später kam heraus, dass die Maschine nicht in der Lage war, selbsttätig Schach zu spielen, sondern dass ein menschlicher Schachspieler in ihr versteckt war, der den Mechanismus bediente. (Wenn heute das Wort »getürkt« als Synonym für »getäuscht« verwendet wird, rührt das wahrscheinlich daher.)

Als im Jahr 1996 der IBM-Computer »Deep Blue« erstmals einen regierenden Schachweltmeister, nämlich Garri Kasparow, in einer Partie mit regulären Zeitkontrollen schlagen konnte, witterte indes niemand einen Betrug. Die Menschheit hatte in den dazwischenliegenden zwei Jahrhunderten den Aufstieg der Technik erlebt und sich an den Gedanken gewöhnt, dass Maschinen so manche Aufgabe übernehmen können und einige Dinge sogar besser bewältigen als der Mensch. Freilich war an »Deep Blue« dennoch etwas Neues, für viele Menschen Verstörendes: Schachspielen galt als äußerst schwieriges Unterfangen, Schachweltmeister wurden zu den klügsten Köpfen der Welt gezählt. Heute hingegen

MENSCH UND MASCHINE In der Barockzeit entwickelte sich ausgehend von René Descartes eine Denkströmung, in der die ganze Welt als eine große Maschine angesehen wurde. Auch Tiere wurden zu Maschinen »degradiert«, und manche Denker bezogen sogar den Menschen in dieses Konzept mit ein. Das populäre Bild von Zahnrädern, die im Gehirn rattern, geht auf diese Vorstellung zurück. ✘

The History of Artificial Intelligence

Machines simulating human behavior have fascinated people since the Baroque age. Artificial intelligence in the true sense of the word emerged soon after World War I; the technology received its name in 1956. In the 1970s, after years of unbridled optimism, research fell into a twenty-year period of “AI winter”—before reemerging all the more powerful.

It was the spring of 1770. The Viennese court was carried away with excitement and enthusiasm after the Austro-Hungarian court official and engineer Wolfgang von Kempelen had presented his most recent work to Empress Maria Theresa: a chess table at which sat a humanoid figure able to move the chess pieces with its mechanical arm—and win against most of its opponents. As the figure was dressed in Ottoman robes, the **AUTOMATON** was soon referred to as the “Turk.” For a long time, people had been racking their brains about how the machine, in whose interior gears and cogs worked away audibly, managed to accomplish this feat. Its inventor, who would keep the mechanism a secret throughout his life, gained fame and honors on his extensive tours of Europe. Many observers suspected deceit: How should some clockwork come close to the intellectual

achievements of humans—at a time when even animals were thought to be mere machines? On the other hand, early proponents of the Enlightenment like the French philosopher Julien Offray de La Mettrie had by then given birth to the idea that humans

MAN AND MACHINE In the Baroque period, René Descartes started a philosophical school of thought in which the entire world was seen as one big machine. Animals were reduced to mere machines, too, and some thinkers even included humans in the concept. The popular image of the cogwheels of the mind goes back to those ideas. ✕

were nothing but machines. Therefore, many a contemporary was convinced that Kempelen’s apparatus was indeed genuine.

In any case, decades later it turned out that his machine was not capable of playing chess by itself, but that a human chess player operating the mechanism was hidden inside. (The German word “getürkt,” which is still used today for “faked,” probably goes back to this episode.)

However, when IBM’s computer “Deep Blue” was the first to defeat a reigning world chess champion, namely Garry Kasparov, in a game under standard time controls in 1996, no one believed in a swindle. In the two centuries that had elapsed, humankind had witnessed the triumphant rise of technology and got used to the idea that machines were able to take over many a task and cope with certain things even better than humans. Admittedly, there was a new dimension to “Deep Blue” many people found irritating: playing chess had previously been classified as an extremely demanding endeavor, and chess players had been considered the world’s greatest masterminds. Today, however, playing chess is rated as a compara-

gilt Schachspielen in der Informatik als vergleichsweise triviale Aufgabe. Dank der modernen Methoden der künstlichen Intelligenz (KI) können auch viel härtere kognitive Nüsse geknackt werden.

Wie Science-Fiction real wurde

Rund 80 Jahre zuvor hätte sich das noch kein Wissenschaftler oder Ingenieur träumen lassen. Damals waren etwa unabhängig vom Menschen handelnde Roboter ausschließlich Sache der Science-Fiction-Literatur und einiger utopischer Filme. Der erste wesentliche Schritt, der das veränderte, kam von dem britischen Logiker und Mathematiker Alan Turing. Im Jahr 1936 publizierte er seine Überlegungen für ein Gerät, das in der Lage sein sollte, »jedes vorstellbare mathematische Problem zu lösen, sofern dieses auch durch einen Algorithmus gelöst werden kann«. Dieser bald als »Turing-Maschine« bezeichnete Mechanismus existierte damals nur auf dem Papier – es gab noch keine technischen Verfahren, um einen solchen realisieren zu können. Die Idee war aber ein entscheidender Anstoß für die Entwicklung der ersten Computer und deren Programmierung. Turing selbst war an der Entwicklung universeller Rechenmaschinen beteiligt: Während des Zweiten Weltkrieges schaffte er es mithilfe elektromechanischer Rechenmaschinen, den Verschlüsselungscode, den Nazi-Deutschland für Nachrichten benutzte (»Enigma«), zu knacken.

Die eskalierenden Kriegshandlungen führten weltweit, auch und vor allem in den USA, dazu, dass mit Hochdruck an der Entwicklung automatisierter Berechnungsverfahren gearbeitet wurde. IBM stellte 1935 die Lochkartenmaschine IBM 601 vor, in Deutschland meldete Konrad Zuse 1937 mehrere Patente für Rechenmaschinen an, in den frühen 1940er-Jahren wurden auch in den USA Digitalrechner entwickelt. Der technische Durchbruch gelang aber erst 1947 durch die Erfindung des Transistors. Fast gleichzeitig lieferte der österreichisch-ungarische, später in den USA tätige Mathematiker John von Neumann die theoretische Basis für Computer.

Eine Motte auf drei Rädern

An diesem technologischen Wettlauf beteiligte sich auch der US-Mathematiker Norbert Wiener: 1940 hatte er einen Algorithmus zur Lösung partieller Differentialgleichungen entwickelt; dann suchte er im Auftrag der US-Armee nach einem Raketenabwehrsystem. Nach Kriegsende heuerte er am Massachusetts Institute of Technology (MIT) an, wo er 1946 die Grundlagen für eine neue Wissenschaft, die **KYBERNETIK**, schuf. Angestachelt durch ein Gerät namens »Homöostat«, das 1946 von seinem britischen Forscherkollegen W. Ross Ashby gebaut und als erste »denkende Maschine« gepriesen wurde, konstruierte Wiener 1949 einen »neurotischen« Roboter: »Palomilla«. Dieses knapp einen halben Meter große dreirädrige Gerät hatte an der Vorderseite zwei Solarzellen, die das lenkbare Vorderrad so steuerten, dass »Palomilla« immer einer Lichtquelle entgegenstrebte – wie eine Motte.

KYBERNETIK In den 1940er-Jahren entwickelte Norbert Wiener eine Wissenschaft der Steuerung und Regelung von Maschinen, lebenden Organismen und sozialen Organisationen und nannte diese »Kybernetik« (griech. Steuermannskunst). Sie ist eine der Wurzeln künstlicher Intelligenz. Die Wissenschaftsrichtung lief sich in den 1960er-Jahren tot, der Name lebt aber im Begriff »cyber« weiter. ✕

tively trivial task in computer science. Thanks to the modern methods of artificial intelligence (AI), it is possible to crack even tougher cognitive nuts.

How science fiction became a reality

About eighty years ago, no scientist or engineer would have dreamt of it. Then, robots acting independently of humans were exclusively a matter of science fiction literature and some utopian films. The first crucial step that changed everything was undertaken by the British logician and mathematician Alan Turing. In 1936, Turing published his ideas for a device that was supposed to be capable of “performing any conceivable mathematical computation if it were representable as an algorithm.” This mechanism, which soon became known as the “Turing Machine,” then only existed on paper—a technical method to realize it did not yet exist. However, the idea gave decisive impetus to the development of the first computer and how it was programmed. Turing himself was involved in the development of universal computing machines; in World War II, using electromechanical calculating devices, he succeeded in cracking the crypto-system used by Nazi Germany (“Enigma”) to encode its messages.

With warfare escalating, efforts to advance automated computing methods were intensified around the globe, also and particularly in the United States. In 1935, IBM presented its punched-card machine IBM 601; in Germany, Konrad Zuse filed multiple patent applications for computing machinery in 1937; in the early 1940s, digital computers were also developed in the United States. However, the technological breakthrough was not accomplished until 1947, when the transistor was invented. Around the same time, the Austro-Hungarian mathematician John von Neumann, who later worked in the United States, supplied the theoretical foundations for computers.

A moth on three wheels

The technology race was also joined by US mathematician Norbert Wiener. By 1940 he had developed an algorithm to solve partial differential equations; after that, he went in search of a missile defense system on behalf of the US army. When the war was over, he applied at the Massachusetts Institute of Technology (MIT), where he laid the basis for a new

science called **CYBERNETICS** in 1946. In 1949, spurred on by a device called “Homeostat,” which had been built by his British fellow researcher W. Ross Ashby and was praised as the first “thinking machine,” Wiener constructed a “neurotic” robot

named “Palomilla”: a tricycle cart not even fifty centimeters in length, it was furnished with two solar cells at the front that directed the steerable front wheel in such a way that “Palomilla” would always be heading for a light source—like a moth.

CYBERNETICS In the 1940s, Norbert Wiener developed a science of the control and regulation of machines, living organisms, and social organizations, which he called “cybernetics” (ancient Greek for “helmsmanship”). It is one of the roots of artificial intelligence. As a science, it had run its course in the 1960s, but its name has lived on in the word component “cyber.” ✦

Die genannte Erfindung war explizit mit der Frage verknüpft, ob und wie sich das Verhalten von Lebewesen mit technischen Mitteln realisieren ließe. So wurde die 200 Jahre alte Idee vom »Menschen als Maschine« wieder virulent: Man hatte aber nicht mehr ein Uhrwerk mit Zahnrädern vor Augen, vielmehr rückte der Begriff »Information« in den Mittelpunkt. 1944 hatte der österreichische Nobelpreisträger für Physik Erwin Schrödinger – der am Alpbacher Ortsfriedhof seine letzte Ruhestätte fand – das Buch *Was ist Leben?* veröffentlicht, in dem er Vererbung mit der Weitergabe von Information gleichsetzte. Das Medium dieses Prozesses, nämlich die DNA als Informationsträger, war damals noch nicht entdeckt. Aber diese Idee führte zu der Theorie, dass alle Lebensprozesse – von der Vererbung bis hin zu kognitiven Fähigkeiten – in ihrem Kern Informationsverarbeitungsprozesse sind. Von hier war es nicht mehr weit zu der Idee, diese in Maschinen nachzubilden. Die brennende Frage dieser Jahre war also: Ist Denken außerhalb des menschlichen Kopfes möglich?

Alan Turing stellt die Frage, ob Maschinen denken können

Es war einmal mehr Alan Turing, der das Problem als Erster klar formulierte: Im Jahr 1950 erschien sein Aufsatz »Computing Machinery and Intelligence«, in dem er explizit die Frage stellte: »Können Maschinen denken?« Mit einer Antwort konnte er nicht dienen – und zwar mit dem Hinweis, dass der Begriff »Denken« nur schwer zu definieren sei. Er dachte sich aber mit dem »Turing-Test« eine Methode aus, mit der man grundsätzlich feststellen könne, ob eine Maschine klug ist: Ein Computer sollte demnach dann als intelligent gelten, wenn er sich über eine elektronische Verbindung mit einem Menschen unterhalten konnte und der Mensch am anderen Ende nicht wusste, ob er sich mit einem anderen Menschen oder mit einer Maschine unterhielt.

Wie wichtig seine Ideen schon bald werden sollten, konnte Turing selbst nicht mehr erleben. Er starb 1954, wahrscheinlich durch Suizid. Der entscheidende Punkt, an dem die Saat der Pioniere aufzugehen begann, kam im Sommer des Jahres 1956: Eine Forschergruppe um John McCarthy und Marvin Minsky lud einige Kollegen zum **DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE** ans Dartmouth College in Hanover, New Hampshire, ein – hier wurde erstmals explizit der Begriff »künstliche Intelligenz« verwendet.

DARTMOUTH-KONFERENZ John McCarthy, Marvin Minsky, Claude Shannon und sieben weitere Forscher nahmen im Sommer 1956 an einer Konferenz am Dartmouth College in Hanover, New Hampshire, teil, die als Startschuss für die KI-Forschung gilt: Im Antrag für Forschungsmittel der Rockefeller-Stiftung wurde erstmals der Begriff »künstliche Intelligenz« verwendet. ✘

Die Geburtsstunde der KI

Die Ziele dieser Konferenz wurden im Förderantrag an die Rockefeller Foundation sehr klar dargelegt: »Wir schlagen vor, im Laufe des Sommers 1956 über zwei Monate ein Seminar zur künstlichen Intelligenz mit zehn Teilnehmern am Dartmouth College durchzuführen. [...] Die Studie soll von der Annahme ausgehen, dass grundsätzlich alle Aspekte des Lernens und anderer Merkmale der Intelligenz so genau beschrieben werden können, dass eine Maschine zur Simulation dieser Vorgänge gebaut werden kann. Es soll versucht

Said invention was explicitly linked to the question whether and how the behavior of living beings could be mimicked with the means of technology. It was then that the two-hundred-year-old idea of “man as a machine” resurfaced as a topic: however, scientists no longer had to make do with some clockwork operated by a set of gears, but rather focused on the term “information.” In 1944, the Nobel-Prize-winning Austrian physicist Erwin Schrödinger—who, by the way, found his final resting place on the Alpbach village cemetery—published his book *What Is Life?*, in which heredity is equated with the transmission of information. The medium of this process, namely DNA as the carrier of information, had not been discovered by then. But the idea led to the theory that all vital processes—from inheritance to cognitive abilities—eventually boil down to information processing. This was not far from the idea of mimicking it with machines. The burning question of those years was thus: Is thinking outside the human brain possible?

Alan Turing wonders whether machines can think

Once again, Alan Turing was the first to clearly formulate the problem; in his essay “Computing Machinery and Intelligence,” which appeared in 1950, he explicitly asked: “Can machines think?” He was not able to supply

an answer, though—conceding that it was difficult to define the term “think.” However, by inventing the Turing test, he contrived a method to basically determine whether a machine was smart. A computer was considered intelligent when it was able to communicate with a person via

DARTMOUTH WORKSHOP John McCarthy, Marvin Minsky, Claude Shannon, and seven other researchers took part in a two-month workshop at Dartmouth College in Hanover, New Hampshire, in the summer of 1956, which is considered today as the kickoff of AI research. In fact, it was in the funding proposal submitted to the Rockefeller Foundation that the term “artificial intelligence” was first used. ✦

electronic connections and the person at the other end was unable to tell whether he or she had been conversing with another human being or with a machine.

Turing did not live to see how seminal his ideas would soon become. He died in 1954, presumably by committing suicide. The crucial moment when the pioneers’ seeds bore fruit followed in the summer of 1956. A group of researchers rallying around John McCarthy and Marvin Minsky invited several colleagues to take part in the **DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE** at Dartmouth College in Hanover, New Hampshire—which was when the expression “artificial intelligence” was explicitly used for the first time.

The birth of AI

The goals of this conference were pointed out in the application for funds submitted to the Rockefeller Foundation: “We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College... The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines

werden herauszufinden, wie Maschinen dazu gebracht werden können, Sprache zu benutzen, Abstraktionen vorzunehmen und Konzepte zu entwickeln, Probleme von der Art, die zurzeit dem Menschen vorbehalten sind, zu lösen, und sich selbst weiter zu verbessern.«

Jeder der Teilnehmer sollte zunächst seine bisherigen Gedanken zu diesen Themenbereichen schriftlich zusammenfassen und an die anderen Teilnehmer schicken. Die Konferenz selbst wurde dann zu einer Art Brainstorming, das nur ein Monat dauerte und ohne unmittelbare Ergebnisse wie etwa Protokolle oder Publikationen zu Ende ging. Allerdings: Die Teilnehmer, die damals meist noch am Beginn ihrer akademischen Laufbahn standen, begannen in der Folge mit Hochdruck an diesen Themen zu arbeiten und wurden praktisch alle zu renommierten Fachleuten auf dem – neuen – Gebiet der KI.

Aufbruchsstimmung und viele Versprechungen

Die nächsten Jahre waren von einer immensen Aufbruchsstimmung geprägt. Die Entwicklung verlief rasant: Auf der einen Seite machte die Computertechnik riesige Fortschritte – Anfang der 1960er-Jahre kamen integrierte Schaltkreise bzw. Computerchips auf den Markt. Auf der anderen Seite gab es viele theoretische Fortschritte. McCarthy beispielsweise entwickelte die Programmiersprache LISP, die rasch zur wichtigsten Plattform für KI-Systeme wurde. Schon bei der Dartmouth-Konferenz waren die Grundzüge des Programms »Logic Theorist« präsentiert worden, das nicht nur Zahlen, sondern auch Symbole und Zeichen verarbeiten konnte. Daraus entstanden die Grundlagen für die Verarbeitung von Sprache und für das Erkennen von Sinnzusammenhängen. Zu einem weiteren großen Thema wurde schon damals die Nachbildung der Funktionsweise des Gehirns in Form von künstlichen neuronalen Netzen: Die wichtigsten Theorien dafür wurden in den 1960er-Jahren entwickelt – und sie gelten im Wesentlichen noch heute. Mit der Idee neuronaler Netze löste man sich erstmals von einem ausschließlich algorithmischen Zugang zu Problemen.

Mit solchen Ansätzen wurden an zahlreichen Universitäten große KI-Forschungszentren aufgebaut, etwa an der Carnegie Mellon University in Pittsburgh, Pennsylvania, am MIT in Cambridge, Massachusetts, und an der Stanford University in Kalifornien. Auch das Militär und die Industrie investierten viel Geld in die Forschung auf diesem Gebiet. Nicht nur in den USA, sondern auch in Europa und in Japan wurden spezifische Förderprogramme gestartet.

Anfangs gab es viele Erfolgsmeldungen: 1959 wurde beispielsweise ein Programm für das Brettspiel Dame vorgestellt, das es mit den besten menschlichen Spielern aufnehmen konnte. Dabei wurde erstmals selbstständiges Lernen von Maschinen realisiert: Das Programm wurde nämlich darauf angesetzt, gegen sich selbst zu spielen und dabei die Wahrscheinlichkeit aufzuzeichnen, mit der ein bestimmter Zug in einer gewissen Spielsituation die Chancen auf den Sieg erhöht. In den frühen 1960er-Jahren blickte die Forschungswelt gebannt auf das Programm **GENERAL PROBLEM SOLVER**, das den Anspruch hatte, mit

GENERAL PROBLEM SOLVER 1957 wurde die Arbeit am »General Problem Solver« begonnen, der eine allgemeine Problemlösungsmethode bieten sollte: Ein Problem sollte dadurch gelöst werden, dass man es in Teilprobleme zerlegte, diese löste und die Lösungen wieder zusammensetzte. Ende der 1960er-Jahre wurde dieser Ansatz als gescheitert fallengelassen. ✘

use language, form abstraction and concepts, solve kinds of problems now reserved for humans, and improve themselves.”

Participants were asked to sum up their earlier thoughts and ideas on these subject areas in writing and send them to the other participants. The conference itself turned into a kind of brainstorming that eventually only lasted for a period of one month and ended without any immediate results, such as records or publications. Yet the participants, most of whom were at the beginning of their academic careers, began to concentrate on these themes in their work: practically all of them became renowned experts in the—new—domain of AI.

A spirit of optimism and lots of promises

The coming years were dominated by a spirit of huge optimism. Developments advanced rapidly. On the one hand, computer technology had made a huge step forward: integrated circuits or computer chips were brought to the market in the early 1960s. On the other hand, theory progressed in many aspects. McCarthy, for example, developed the programming language LISP, which soon became one of the most important platforms for AI systems. The basic principles of the program “Logic Theorist,” which was able to not only work with numbers but also with signs and symbols, had already been presented at the Dartmouth Workshop. This eventually provided the basis for processing language and recognizing complex meanings. Even then, simulating the function of the human brain in the form of artificial neural networks had become another major theme: the most seminal theories were developed in the 1960s—and are essentially still valid today. The idea of neural networks made it possible for the first time to break away from a purely algorithmic approach to problem solving.

Big AI research centers focusing on these methods were installed at numerous universities, including Carnegie Mellon University in Pittsburgh, Pennsylvania, MIT in Cambridge, Massachusetts, and Stanford University in California. The military and industries likewise invested a lot of money in research in those fields. Specific support programs were launched not only in the United States, but also in Europe and Japan.

In the beginning, there were many success stories. For example, in 1959, a program for the board game checkers was introduced that could compete with the best human players.

GENERAL PROBLEM SOLVER In 1957, work began on the General Problem Solver, which was intended to provide a universal problem-solving method. A problem was supposed to be solved by splitting it up into partial problems, solving those, and recombining the individual solutions. In the late 1980s, the approach was eventually considered failed and was dropped. ✦

This was the first time a machine had been realized that was capable of learning autonomously: the program had been devised in such a way that it would play against itself whilst recording the probability with which a

particular move in a particular situation during the game would improve the chances to win. In the 1960s, researchers were spellbound by the program **GENERAL PROBLEM SOLVER**, which claimed that it could solve any problem with simple methods—namely by breaking a problem down into subproblems that could be tackled separately.

einfachen Methoden beliebige Probleme lösen zu können – und zwar dadurch, dass es ein Problem in seine Teilprobleme zerlegte, die dann separat gelöst werden konnten.

Für großes Aufsehen sorgte der deutschstämmige Informatiker Joseph Weizenbaum, der 1966 am MIT den ersten Chatbot namens »ELIZA« präsentierte. Dabei wurde ein Dialog zwischen einem Psychotherapeuten und einem Patienten simuliert. Das Überraschende war – auch für den Forscher selbst –, dass man auf relativ einfache Weise Menschen die Illusion eines beseelten Gesprächspartners vermitteln konnte. In Stanford wurde um diese Zeit mit der Entwicklung des ersten teilautonomen Roboters begonnen: »Shakey« konnte sich mit Kameras und Sensoren ein Bild seiner Umgebung machen, konnte gesprochene Sprache erkennen und stand per Funk mit einem Zentralrechner in Verbindung.

Zu große Versprechungen: Der Anbruch des »KI-Winters«

Diese Jahre waren durch eine fast grenzenlose Erwartungshaltung im Hinblick auf die Fähigkeit von Computern geprägt, was sich auch in der Populärkultur dieser Zeit niederschlug (man denke etwa an *Raumschiff Enterprise*). Allerdings wurde immer deutlicher, dass die tatsächlichen Erfolge bei Weitem nicht an die vollmundigen Versprechungen herankamen. Man hatte unterschätzt, wie komplex Denken und Sprache sind. Zudem blieben trotz aller Verbesserungen die Rechen- und Speicherkapazitäten der Hardware hinter der erforderlichen Leistung zurück. Und es fehlte, wie die spätere Entwicklung zeigen sollte, an ausreichend großen und validen Datenbanken mit Lerndaten für das maschinelle Lernen; vor 50 Jahren waren zum Beispiel nicht einmal die großen Enzyklopädien digitalisiert.

Die Folge war, dass die Euphorie Ende der 1960er-Jahre in sich zusammenbrach. Die Politik kürzte die Fördermittel, die Industrie reduzierte ihre Investitionen in visionäre Forschungsprojekte. Es begann der »Winter der künstlichen Intelligenz«, der zwei Jahrzehnte andauern sollte.

Viele kleinere, konkretere Schritte

In dieser Zeit gab es kaum visionäre Grundlagenforschung in Sachen KI. Die Industrie konzentrierte sich auf konkrete Projekte zur Weiterentwicklung der Hard- und Software. Bei der Entwicklung von Anwendungen wurden freilich die erzielten Errungenschaften genutzt. Die Arbeiten am »General Problem Solver« waren zwar eingestellt worden, daraus entwickelten sich aber in den 1970er-Jahren die ersten praxistauglichen **EXPERTEN-**

SYSTEME wie etwa MYCIN, das mit einigem Erfolg zur Unterstützung von Diagnose- und Therapieentscheidungen bei bestimmten Krankheiten eingesetzt wurde. Ein weiteres Beispiel: 1982 gelang es mithilfe des Programms »Dragon« (später »Nuance«) erstmals, einer Maschine beizubringen, gesprochene Sprache zu verschriftlichen. Man erhoffte sich aus solchen Entwicklungen eine Reihe industrieller Anwendungen, doch auch hier konnten viele der allzu optimistischen Prognosen nicht realisiert werden.

EXPERTENSYSTEME Ein Expertensystem ist ein Computerprogramm, das Menschen bei der Lösung von komplexeren Problemen wie ein Experte unterstützen kann, indem es Handlungsempfehlungen aus einer Wissensbasis ableitet. Darin wird über Wenn-Dann-Beziehungen menschliches Wissen für Computer verständlich dargestellt. ✘

In 1966, the computer scientist Joseph Weizenbaum, a native German, caused a sensation when he presented “ELIZA,” the first chatbot, at MIT, simulating a dialogue between a psychotherapist and a patient. What was so amazing—even for the scientist himself—was that it seemed relatively easy to create the illusion of a living interlocutor in people. At the same time, development of the first semi-autonomous robot began in Stanford: “Shakey” was able to perceive its surroundings thanks to cameras and sensors and recognize spoken language whilst connected to a mainframe computer via a radio link.

Inflated promises: the advent of “AI winter”

Those years were marked by virtually unbridled expectations as to the abilities of computers, also manifesting themselves in contemporary popular culture (just think of *Star Trek*). However, it gradually became apparent that the actual successes fell short of such flamboyant promises by far. The complexity of thought and language had been underestimated. Despite substantial improvements, the computing and memory capacities of hardware lagged behind the performance required. As future developments would show, there was also a lack of sufficiently large and valid databases supplying data for machine learning; fifty years ago the big encyclopedias had not even been digitized, for instance.

All of this resulted in euphoria collapsing in the late 1960s. Governments cut down on their funding, and industries reduced their investments in visionary research projects. The “winter of artificial intelligence,” which would last for two decades, began.

Many smaller and more concrete steps

Hardly any visionary basic research into AI was conducted in those years. Industries concentrated on concrete projects dedicated to the improvement of hardware and software. Earlier achievements were of course made use of in the development of applications. Although work on the “General Problem Solver” had come to a halt, the first feasible **EXPERT**

EXPERT SYSTEMS An expert system is a computer program that can assist humans in solving complex problems like a human expert by reasoning through and deducing recommendations for action from a body of knowledge. Human knowledge is represented as sets of computer-comprehensible if-then rules. ✕

SYSTEMS originated from it in the 1970s, such as MYCIN, which was more or less successfully employed to support decisions related to diagnosis and therapy in the case of certain diseases. Another example: In 1982, it became

possible for the first time to successfully teach a machine how to transcribe spoken language with the aid of the program “Dragon” (later known as “Nuance”). It was hoped that such developments would lend themselves to a number of industrial applications, but many of these all too optimistic prognoses again failed to materialize.

Ein neuer Frühling bricht an

Es dauerte, aber dann ging der »KI-Winter« doch zu Ende – und zwar letztlich infolge des Aufstiegs des Internets. 1993 kam mit »Netscape« der erste für Konsumenten geeignete Browser auf den Markt. Das war der Startschuss für eine bis dahin nicht vorstellbare Vernetzung von Computern. In Kombination mit der rapide gestiegenen Leistungsfähigkeit der Hardware entstanden in der Folge riesige Datenmengen – also genau das, was den KI-Pionieren in den 1960er-Jahren gefehlt hatte, um der Idee des maschinellen Lernens Leben einzuhauchen.

Es ist eine Ironie der Geschichte, dass ausgerechnet jenes Ereignis, das die KI wieder ins Licht der Weltöffentlichkeit rückte, eigentlich noch nichts damit zu tun hatte: Der Supercomputer »Deep Blue«, der 1996 den amtierenden Schachweltmeister in einer Partie besiegte, war »nur« eine äußerst schnell rechnende Maschine, die streng algorithmisch alle möglichen Schachzüge durchprobierte und jene mit der höchsten Siegeschance auswählte. Von wirklich »intelligentem« Verhalten konnte bei dieser rein mechanischen Vorgangsweise also keine Rede sein. Dennoch war die Symbolwirkung gewaltig: Schon im Jahr 1957 hatten die Väter der KI prognostiziert, dass innerhalb der nächsten zehn Jahre ein Computer Schachweltmeister werden würde – eine Zusage, die nicht eingehalten werden konnte. Im Jahr 1990 wurde die Prognose wiederholt, und nun klappte es auf einmal: Das alte Versprechen konnte eingelöst werden – 40 Jahre später.

»Watson« setzte neue Maßstäbe

Dadurch entstand eine neue Aufbruchsstimmung: Big Data, verbesserte Algorithmen und die weiter steigende Rechenleistung ermöglichten einen neuen Forschungs- und Entwicklungsboom, dessen vorläufiger Gipfelpunkt im Jahr 2011 erreicht wurde: Damals gewann »Watson«, ein Nachfolger von »Deep Blue«, die us-Quizshow »Jeopardy!«. Die Aufgabe dabei war, zu einer vorgegebenen Antwort die richtige Frage zu finden. Es reichte also nicht aus, in Blitzesschnelle Datenbanken zu durchforsten oder mögliche Szenarien durchzurechnen, vielmehr mussten Sinnzusammenhänge in der Sprache analysiert und »verstanden« werden. Das gelang »Watson« durch ein selbstlernendes System, das mit Beispieldaten gefüttert wurde und daraus Schlüsse zog. Diese wurden mit wissensbasierten Systemen gekoppelt, aus denen per Inferenzmaschinen neue Aussagen abgeleitet wurden.

Das maschinelle Lernen wurde seither in vielen Anwendungsfeldern zu einer Kerntechnologie: bei der Vernetzung von Produktionsabläufen und der Steuerung von Robotern (Industrie 4.0), bei autonomen Fahrzeugen, bei der Spracherkennung und -steuerung in Smartphones, bei intelligenten Lautsprechern, bei Assistenzsystemen etwa in der Radiologie oder der Krebsforschung, bei Systemen in Anwaltskanzleien, die Fehler in Vertragsentwürfen finden, bei Übersetzungsprogrammen und so weiter und so fort.

Maschinen mit dem Anschein von Kreativität

Auf äußerst publicityträchtige Weise wurde die Macht dieser Methoden erneut bei Spielen demonstriert: 2015 war das Schachprogramm »Giraffe« in der Lage, in 72 Stunden Schach auf Meisterniveau zu lernen. 2016 besiegte

The dawn of a new spring

Although it took a long time, the “AI winter” eventually did come to an end—not least thanks to the rise of the Internet. In 1993 “Netscape” was the first browser to appear on the market that was suitable for consumers. This kicked off a yet unimaginable internetworking of computers and, in conjunction with the rapid increase of hardware capacity, led to the accumulation of huge amounts of data—exactly what the AI pioneers had been lacking in the 1960s in order to breathe life into their idea of machine learning.

It is an irony of history that, of all things, an event that had actually nothing to do with it contributed to AI’s resurfacing in public awareness around the globe. The supercomputer “Deep Blue,” which defeated the reigning world chess champion in a game in 1996, was “merely” an extremely fast computer that tried out all of the possible chess moves strictly algorithmically, choosing those with the best chances to win. This purely mechanical method could not be farther away from truly “intelligent” behavior. Nevertheless, the event made a huge symbolic impact: As early as 1957, the fathers of AI had prophesied that a computer would become world chess champion within the following ten years—a pledge that could not be redeemed. The prognosis was repeated in 1990, and all of a sudden it worked out: the old promise was finally kept—with a forty-year delay.

“Watson” sets new benchmarks

This resulted in a new euphoria: big data, improved algorithms, and constantly increasing computing capacities facilitated a new research and development boom, which, for the time being, reached its peak in 2011, when “Watson,” the successor of “Deep Blue,” won the US quiz show “Jeopardy!”. The task was to find the right question to a given answer. It therefore did not suffice to browse through the contents of databases at cyberspeed and calculate possible scenarios; rather, it was necessary to analyze and “comprehend” the context and meaning of language. “Watson” managed to do so thanks to a self-learning system that had been fed with sample data on the basis of which it drew its conclusions. These were coupled with knowledge-based systems from which new statements were deduced by means of inference machines.

Since then, machine learning has become a core technology applied in countless fields, ranging from the interconnection of production processes and the control of robots (Industry 4.0) to autonomous vehicles; from language recognition and control in smart phones to intelligent speakers and translation software; from assistant systems used in radiology and cancer research to law office systems identifying mistakes in draft agreements, etc.

Machines feigning creativity

The powerful impact of these methods was once again demonstrated in a highly publicity-gaining fashion in the field of games development. In 2015, the chess software “Giraffe” managed to learn how to play chess on a

»AlphaGo« den besten menschlichen Go-Spieler, 2017 zwang »Liberatus« die besten Pokerspieler in die Knie. Diese Systeme verknüpften maschinelles Lernen und Statistik auf eine Weise, die den Anschein erweckten, dass Maschinen kreative Spielzüge entwickelten. Apropos Kreativität: 2017 wurde eine KI-Anwendung erdacht, um die Handlung des Fantasy-Epos *Games of Thrones* von George R. R. Martin fortzuschreiben. Dazu wurde der Computer mit den Inhalten der ersten fünf Bände gespeist. Das lernfähige Programm konnte sich Stil, besondere Ausdrucksweisen und vom Autor gerne verwendete Redewendungen aneignen und dank KI die einzelnen Charaktere fortentwickeln und sogar eine neue Figur erfinden. Ebenfalls im Sommer des Vorjahrs wurde von Forschern in Chicago eine Software vorgestellt, die eigenständig Rezensionen verfassen kann. Eine versuchsweise Beurteilung durch 600 Personen ergab, dass diese die maschinengenerierten Rezensionen im Durchschnitt als ähnlich nützlich ansahen wie von leibhaftigen Menschen verfasste Besprechungen. Fast gleichzeitig präsentierten Forscher der Rutgers Universität ein System, das Bilder malte. Diese Bilder wurden mit von Künstlern erstellten Gemälden vermischt und 18 Experten in einem Blindtest zur Beurteilung vorgelegt: Die Jury beurteilte die von der KI gefertigten Gemälde insgesamt als besser als die von den Künstlern gemalten. In diesen speziellen Fällen haben die KI-Systeme also den »Turing-Test« bestanden.

In Sachen allgemeiner Intelligenz liegen KI-Systeme allerdings weiterhin meilenweit hinter den Fähigkeiten der Menschen zurück. Das zeigt sich etwa beim Loebner-Preis, der seit dem Jahr 1991 alljährlich ausgeschrieben wird. Dabei handelt es sich quasi um eine Weltmeisterschaft im »Turing-Test«: Für die Silbermedaille und ein Preisgeld von 25.000 Euro muss ein Computersystem in einer fünfminütigen schriftlichen Unterhaltung 30 Prozent der Juroren davon überzeugen, dass es ein Mensch ist. Dieser Preis wurde noch niemals vergeben – geschweige denn die Goldmedaille, bei der ein KI-System 25 Minuten lang in verschiedenen Medien – Sprache, Musik, Bild, Video – menschenähnliches Verhalten an den Tag legen muss.

Rückkehr der alten philosophischen Fragen

Trotzdem: Mit dem offenkundigen Erfolg der neuen KI-Systeme kehren die alten philosophischen und gesellschaftspolitischen Fragen, die sich schon die Gründerväter der KI gestellt haben, mit voller Wucht zurück: Können Maschinen denken? Ist Denken unabhängig von der Plattform, auf der die Prozesse ablaufen (biologisch oder technisch)? Ist es möglich, dass sich Maschinen selbst verbessern (was Google laut eigenen Angaben bereits mit dem System »AutomL« gelungen ist)? Müssen wir befürchten, dass uns die Kontrolle über **SUPERINTELLIGENTE** Maschinen entgleitet? Zerstört KI am Ende die Grundlagen unseres Wirtschaftssystems und der Demokratie?

SUPERINTELLIGENZ Mit »Superintelligenz« wird eine Maschine bezeichnet, die über eine dem Menschen in vielen oder allen Gebieten überlegene Intelligenz verfügt. Von diesem Punkt an wird die weitere Entwicklung hauptsächlich von der KI vorangetrieben – und nicht mehr vom Menschen. Mit (technologischer) »Singularität« ist jener Zeitpunkt gemeint, an dem künstliche Intelligenz die menschliche Intelligenz übertrifft. ✖

champion level within 72 hours; in 2016, “AlphaGo” defeated the best human Go player; in 2017, “Liberatus” brought the best poker players to their knees. These systems combine machine learning and statistics in such a way that the machines seem capable of planning their own creative moves. Speaking of creativity: in 2017 an AI application was devised to carry the plot of George R.R. Martin’s fantasy epic *Games of Thrones* further. For this, a computer was fed with the contents of the first five volumes. The self-learning software was able to take in the author’s style, special expressions, and preferred phrases and, thanks to AI, managed to develop the characters further, which even included the introduction of a new figure. Also in the summer of last year, scientists in Chicago presented a software program capable of writing reviews. A trial run during which the system was put to the test revealed that 600 subjects found these machine-generated reviews just as helpful as those penned by human authors. At around the same time, Rutgers University researchers presented a system generating paintings. These pictures were then jumbled with paintings made by real artists and submitted to eighteen experts for blind testing: the jury rated the paintings produced by AI higher than those executed by the artists. In these particular cases, AI can well be said to have passed the Turing test.

However, when it comes to general intelligence, AI systems still lag way behind human abilities. This is proven, for example, by the Loebner Prize, an annual competition that has been organized since 1991—a kind of world championship based on the Turing test. For a silver medal and a cash prize of 25,000 euros, a computer system would have to be able to convince thirty percent of the jury members during a five-minute textual conversation that it is a human being. This prize has never been awarded—let alone the gold medal, for which an AI system would have to simulate human behavior for a period of 25 minutes, in multiple media: language, music, image, and video.

The return of the old philosophical questions

But still: in the wake of the new AI systems and their evident success, the old philosophical and sociopolitical questions that were already posed by the founding fathers of AI are closing in on us with full force: Can machines think? Is thought independent of the platform on which the processes take place (biologically or technically)? Is it possible for machines to improve autonomously (which, by its own account, Google has achieved with its “AutoML” system)? Are we running the risk of losing control over

SUPERINTELLIGENT machines? Will AI eventually destroy the foundations of our economic system and of democracy? The discussion remains open, as do the direction in which AI will continue developing and the economic and social spheres in which the new technologies will be taking hold in the near future. In any case, more

SUPERINTELLIGENCE “Superintelligence” refers to a machine that possesses intelligence surpassing that of humans in many, or all, areas. From this point forward, development is mostly driven by AI—and no longer by humans. (Technological) “singularity” refers to that point in time when artificial intelligence starts exceeding, and moving beyond, human intelligence. ✦

Die Diskussion ist ebenso offen wie die Richtung, in die sich KI weiterentwickeln wird, und in welchen Bereichen von Wirtschaft und Gesellschaft die neuen Technologien wie rasch Platz greifen werden. Immer mehr Staaten und Organisationen beginnen jedenfalls, sich darauf vorzubereiten. Nicht alle gehen dabei so weit wie die Vereinigten Arabischen Emirate, die Ende 2017 als erstes Land der Welt ein eigenes »Ministerium für künstliche Intelligenz« eingerichtet haben. Aber auch Österreich hat seit dem Vorjahr immerhin einen Roboter- rat und seit heuer ein Bundesministerium für Digitalisierung.

KI und Österreich (Helmut Leopold)

KI-Forschung gibt es in **ÖSTERREICH** schon seit den 1980er-Jahren. Bereits 1984 wurde das Österreichische Forschungsinstitut für Artificial Intelligence (OFAI) gegründet. Diesem steht der KI-Pionier Robert Trappl seit der Gründung als Leiter vor, 2012 wurde er zum Präsidenten der International Academy for Systems and Cybernetic Sciences gewählt.

Für den Durchbruch vieler aktueller KI-Anwendungen sorgte eine österreichisch-schweizerische Technologiekompetenz. Bereits 1997 haben Sepp Hochreiter, der Leiter des KI-Labs am Linz Institute of Technology (LIT) der Johannes-Kepler-Universität (JKU), und Jürgen Schmidhuber, der Leiter von IDSIA, einem schweizerischen Forschungsinstitut für künstliche Intelligenz, die grundlegende KI-Technologie LSTM (»long short-term memory«) entwickelt.

Am Austrian Institute of Technology (AIT) konnten erfolgreich mehrere internationale führende KI-Initiativen etabliert werden, die aktuell in nationalen und in EU-Forschungsinitiativen weiterverfolgt werden: Beispiele dafür sind die Analyse von sehr umfangreichen Bild- und Audiodaten zur öffentlichen Sicherheit und zum Schutz kritischer Infrastrukturen, die Analyse von sehr großen Textdateien für Polizei- und Gerichtsakten, Verkehrsanalysen von sehr großen Datenmengen im Internet und in Telekommunikationsnetzwerken, die Analyse von großen heterogenen Datenstrukturen in historischen Dokumenten, die effektive Spracheingabe für modernes Dokumentenmanagement oder die automatische Anomalie-Erkennung in IT-Systemen, mit deren Hilfe frühzeitig komplexe Cyberangriffe selbstständig erkannt werden können. ✘

HOTSPOTS IN ÖSTERREICH Österreich ist ein internationaler Vorreiter bei der Verwendung mathematischer Logik in der KI. Viel Pionierarbeit hat z. B. Georg Gottlob (TU Wien) geleistet. Mit einem ganz besonderen Thema beschäftigt sich Gerhard Widmer an der Universität Linz: Er analysiert Musik und ihre Interpretation mithilfe von KI-Methoden. ✘

Empfehlenswerte Literatur zur Geschichte der KI

- Manuela Lenzen, *Künstliche Intelligenz. Was sie kann & was uns erwartet*, München 2018.
- Thomas Ramge, *Mensch und Maschine. Wie Künstliche Intelligenz und Roboter unser Leben verändern*, Ditzingen 2018.
- Thomas Rid, *Maschinendämmerung. Eine kurze Geschichte der Kybernetik*, Berlin 2016.
- Tom Standage, *Der Türke. Die Geschichte des ersten Schachautomaten und seiner abenteuerlichen Reise um die Welt*, Frankfurt am Main 2002.
- Max Tegmark, *Leben 3.0. Mensch sein im Zeitalter Künstlicher Intelligenz*, Berlin 2017.

and more states and organizations have taken measures to brace themselves for it. Not all of them are going as far as the United Arab Emirates as the world's first government to have installed an "Artificial Intelligence Department" in late 2017. But since last year, Austria has at least had a "Robot Council" to boast of, and a "Federal Ministry for Digitization" has been installed this year.

AI and Austria (Helmut Leopold)

In **AUSTRIA**, research into AI has been undertaken since the 1980s. The Austrian Research Institute for Artificial Intelligence (OFAI) was established as early as 1984. Since its foundation it has been headed by the AI pioneer Robert Trapp, who was elected president of the International Academy for Systems and Cybernetic Sciences in 2012.

Austro-Swiss technology expertise has ensured the breakthrough of many recent AI applications. In 1997, Sepp Hochreiter, head of the AI laboratory at the Linz Institute of Technology (LIT) of the Johannes Kepler University (JKU), and Jürgen Schmidhuber, head of IDSIA, a Swiss research institute for artificial intelligence, developed the groundbreaking AI technology LSTM ("long short-term memory").

Several internationally leading AI initiatives that are currently being looked into within the framework of national and EU research projects have successfully been established at the Austrian Institute of Technology

(AIT). Examples include the analysis of huge mounds of image and audio data for the benefit of public security and the protection of critical infrastructures; the analysis of extensive text files for police and court records; traffic analyses of immense data

volumes on the Internet and in telecommunications networks; the analysis of large heterogeneous data structures in historical documents; an efficient speech input system for state-of-the-art document management; or automatic anomaly detection in IT systems capable of autonomously identifying complex cyberattacks at an early stage. ✕

HOTSPOTS IN AUSTRIA Austria is an international pioneer in using mathematical logic in AI. Seminal work in the field was done, for example, by Georg Gottlob (Vienna University of Technology). A very specific subject is explored by Gerhard Widmer from the University of Linz, Upper Austria. He analyzes music and its interpretation using AI methods. ✕

Recommended literature on the history of AI

- Manuela Lenzen, *Künstliche Intelligenz. Was sie kann & was uns erwartet*, Munich 2018.
- Thomas Ramge, *Mensch und Maschine. Wie Künstliche Intelligenz und Roboter unser Leben verändern*, Ditzingen 2018.
- Thomas Rid, *Maschinendämmerung. Eine kurze Geschichte der Kybernetik*, Berlin 2016.
- Tom Standage, *Der Türke. Die Geschichte des ersten Schachautomaten und seiner abenteuerlichen Reise um die Welt*, Frankfurt am Main 2002.
- Max Tegmark, *Leben 3.0. Mensch sein im Zeitalter Künstlicher Intelligenz*, Berlin 2017.

Die Technik hinter der künstlichen Intelligenz

Bei modernen Methoden der künstlichen Intelligenz werden die herkömmliche mathematische Logik sowie die darauf aufbauenden regelbasierten Algorithmen und statistischen Verfahren durch künstliche neuronale Netze und weitere Ansätze der Entscheidungsfindung ergänzt.

Die Entwicklung von künstlicher Intelligenz (KI), wie der Versuch der Automatisierung intelligenten menschlichen Verhaltens durch lernende Maschinen bezeichnet wird, ist in jüngster Zeit in eine neue Phase des Booms eingetreten, der jedoch keineswegs vom Himmel gefallen ist. Schon 1956, also vor mehr als 60 Jahren, hatte John McCarthy in den USA eine Konferenz unter dem Namen **DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE** abgehalten. Dieser Workshop war Teil eines Forschungsprojektes, mit dem der Begriff KI begründet wurde.

Im Zuge der rasanten Entwicklung der EDV und der Informatik als Wissenschaftsdisziplin ging man der Frage nach, welche kognitiven Fähigkeiten intelligentes Verhalten von Menschen und welche Grundfunktionen intelligente Fähigkeiten von Maschinen bestimmen. Alan Turing hatte bereits 1950 Wissensrepräsentation nicht auf Daten beschränkt,¹ sondern als Fähigkeit definiert, Schlussfolgerungen zu ziehen, und Sprachverständnis und Lernfähigkeit als zentrale Intelligenzkategorien bezeichnet. Schließlich entwickelten sich die Kognitionswissenschaften und die kognitiven Neurowissenschaften, die heute von der modernen Informatik der KI getrennt verfolgt werden.

Obwohl die grundlegende Informatikdisziplin der mathematischen Logik, die auf griechische Philosophen wie Aristoteles und Sokrates zurückgeht, auch eine wichtige Grundlage für die KI-Forschung war und ist, galt eine ihrer zentralen Fragen immer schon dem Aspekt, dass intelligentes Verhalten nicht immer logischen Gesetzen entspricht. Grundlegende Fragestellungen der KI-Forschung sind, wie man mit Unsicherheiten und Widersprüchen umgeht und warum – eine frühe Erkenntnis der Kognitionsforschung – rationales Handeln nicht immer aufgrund einer vorhandenen Datenlage erfolgt.

Die KI-Entwicklung durchlief verschiedene Phasen, mit der die Grenzen maschinellen logischen Denkens, menschenähnlicher Problemlösung und Entscheidungsfindung sowie des Lernens aus Erfahrung in verschiedenen Anwendungsszenarien immer wieder neu ausgelotet wurden.

In unterschiedlichsten Anwendungsgebieten kommen nun mehr und mehr KI-Ansätze zum Einsatz, da mit ihnen immer wieder weitaus bessere Ergebnisse erzielt werden als mit herkömmlichen Lösungsansätzen. Die bekanntesten praktischen Einsätze sind die Mustererkennung von Objekten für Kameras von autonomen Fahrzeugen (ein Bereich, in dem Google einen großen Innovationsvorsprung aufweisen kann), Spracherkennung Alexa

Bei der **Dartmouth**-Konferenz 1956 wurden folgende Themen behandelt:

- _ Automatische Computer
- _ Wie muss ein Computer programmiert werden, um eine Sprache zu benutzen
- _ Neuronale Netzwerke
- _ Theoretische Überlegungen zum Umfang einer Rechenoperation
- _ Selbstverbesserung
- _ Abstraktionen
- _ Zufälligkeit und Kreativität ✕

¹ Daten allein haben keine besondere Bedeutung. Erst durch ihre Verwendung in bestimmten Kontexten werden sie zu Information und schließlich zu Wissen.

Helmut Leopold

The Technology Behind Artificial Intelligence

Modern methods of artificial intelligence supplement the traditional mathematical logic as well as the rule-based algorithms and statistical procedures based on it with artificial neural networks and other decision-finding approaches.

The development of artificial intelligence (AI), i.e., the attempt to automate intelligent human behavior in learning machines, has recently entered into a new boom phase, which, however, comes as no surprise. More than sixty years ago, John McCarthy organized a conference under the name **DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE** in the United States in 1956. The workshop was part of a research project in whose context the term AI was coined.

The rapid development of EDP and computer science as a discipline in its own right raised the question which cognitive skills define human intelligent behavior and which basic functions the intelligent abilities of machines. Alan Turing had already not restricted the representation of knowledge to data¹ but defined it as the capability to draw conclusions and described the ability to understand language and to learn as crucial categories of intelligence. The cognitive sciences and the neurosciences that finally emerged are dealt with as separate disciplines by modern AI computer science.

Although mathematical logic as the fundamental discipline of computer science, which goes back to Greek philosophers like Aristotle and Socrates, was and is an important foundation for AI research, one of its crucial questions has always been concerned with the aspect that intelligent behavior does not always follow logic rules. Basic issues of AI research are how to deal with insecurities and contradictions and to answer the question why rational actions are not always based on the given data (an insight that ranks among the early findings of cognitive science).

The development of AI passed through various phases in which the limits of machines' logical thinking, human-like problem solving, decision-finding, and learning from experience have again and again been sounded anew in different areas of application.

¹ Data as such have no special meaning; it is only their use within certain contexts that turns them into information and, finally, into knowledge.

At the 1956 **Dartmouth** Workshop, the following topics were addressed:

- _ Automatic Computers
- _ How Can a Computer be Programmed to Use a Language
- _ Neuron Nets
- _ Theory of the Size of a Calculation
- _ Self-Improvement
- _ Abstractions
- _ Randomness and Creativity ✖

(Amazon), Siri (Apple), Bixby (Samsung) und Assistant (Google), Googles Übersetzung von Text in alle Sprachen, der Boom zahlreicher Chatbots² für eine neue Art der Kundenschnittstelle bei Hotlines, Onlineverkauf und Supply-Chain-Management sowie KI-Algorithmen, um immer komplizierter werdende IT-Systeme vor ausgefeilten Cyberangriffen zu schützen.³

Chatbots wird diesbezüglich eine hohe wirtschaftliche Bedeutung zukommen. Eine Oracle-Studie geht davon aus, dass bis zum Jahr 2020 rund 80 Prozent der Unternehmen Chatbots in der Kundenbetreuung einsetzen werden.⁴ Gartner bestätigt diese Annahme mit der Aussage, dass bis 2021 mehr als die Hälfte der us-Unternehmen jährlich mehr in Chatbots investieren wird als in Apps.⁵

Leistungsfähige Hardware als Grundlage für KI

Die Durchbrüche der beeindruckenden Leistungsfähigkeit von KI-Algorithmen, die auf computerbasierten neuronalen Netzen (NN) beruhen, gehen im Wesentlichen mit der Verfügbarkeit von sehr leistungsstarken Prozessoren und günstigen Speicherkapazitäten einher; typischerweise müssen für Lernprozesse laufend mehrere Millionen Parameter in kürzester Zeit berechnet werden.

Vor allem die Spieleindustrie mit ihrer Anforderung sehr leistungsstarker Prozessoren zur Bearbeitung von hochauflösenden dynamischen Grafiken und Bildern hat einen wesentlichen Anteil am aktuellen Durchbruch der KI-Technologien. Für Spielkonsolen wurden spezielle Grafikkarten, sogenannte GPUs (Graphic Processing Units), mit Prozessoren entwickelt, die sich durch hohe Leistungsfähigkeit und effiziente Verarbeitung von parallelen Rechenoperationen auszeichnen. Durch die dank des Spielkonsolenmarkts hohen produzierten und verkauften Stückzahlen sind potente Rechenleistungen verfügbar und können wirtschaftlich in verschiedenen Anwendungen eingesetzt werden.

Zudem geht die Entwicklung der leistungsfähigen Prozessoren rasant weiter. Huawei ist mit dem Mate 10 Pro und dem hauseigenen Prozessor Kirin 970 in den internationalen Wettlauf eingestiegen. Huawei nennt seinen Prozessor NPU (Neural Processing Unit); seine Rechenleistung soll stärker und seine Geschwindigkeit größer sein als die aller bisherigen GPUs. Nach Eigenangaben wurden im Kirin 970 rund 5,5 Milliarden Transistoren auf einem Quadratzentimeter verbaut. Den menschlichen Gehirnzellen vergleichbar bedeuten mehr Transistoren mehr Denkleistung, weshalb bestimmte Algorithmen parallel ausgeführt werden können.

Google nennt seine KI-Spezialhardware Tensor Processing Unit (TPU)⁶, IBM hat seinen neuromorphen Prozessor TrueNorth, AMD kaufte bereits 2006 den GPU-Spezialisten ATI⁷, 2016 erwarb Intel die auf die Entwicklung von Deep-Learning-Prozessoren spezialisierten Unternehmen Nervana und Mobileye⁸. Laufend werden in diesem Highend-Prozessorenbereich für unterschiedliche Märkte Start-ups wie KnuEdge, Eyeriss, krtrl, Graphcore, Brain-Chip, TeraDeep, Wave Computing, Horizon Robotics, NeuroAM3, P-Neuro oder Spinnaker aktiv.⁹

² Kunstwort aus »chat« (plaudern) und »bot« (Roboter).

³ Siehe Helmut Leopold, Thomas Bleier und Florian Skopik (Hg.), *Cyber Attack Information System. Erfahrungen und Erkenntnisse aus der IKT-Sicherheitsforschung*, Wiesbaden 2015.

⁴ Siehe *Can Virtual Experiences Replace Reality? The future role for humans in delivering customer experience*, hg. von Oracle und Coleman Parkes, Redwood Shores 2016; https://www.oracle.com/webfolder/s/delivery_production/docs/FY16h1/doc35/CXresearchVirtualExperiences.pdf (20.05.2018).

⁵ Siehe »Gartner Says 25 Percent of Customer Service Operations Will Use Virtual Customer Assistants by 2020«; <https://www.gartner.com/newsroom/id/3858564> (20.05.2018).

⁶ Siehe »An in-depth look at Google's first Tensor Processing Unit (TPU)«; <https://cloud.google.com/blog/big-data/2017/05/an-in-depth-look-at-googles-first-tensor-processing-unit-tpu> (20.05.2018).

⁷ Um 5,4 Milliarden us-Dollar.

⁸ Siehe Nicole Hemsöth, »Intel, Nervana Shed Light on Deep Learning Chip Architecture«, in: *The Next Platform*, 11.01. 2018; <https://www.nextplatform.com/2018/01/11/intel-nervana-shed-light-deep-learning-chip-architecture/> (20.05.2018). Siehe auch »Mobileye. Das mächtige Auge des Autos«, in: *Handelsblatt*, 13.03.2017; <http://www.handelsblatt.com/impresum/nutzungshinweise/blocker/?callback=%2Funternehmen%2Findustrie%2Fmobileye-das-maechtige-auge-des-autos%2F19508880.html> (20.05.2018).

⁹ Siehe Shan Tang, »A List of Chip/IP for Deep Learning«, in: *Medium*, 11.08.2017; <https://medium.com/@shan.tang/g/a-list-of-chip-ip-for-deep-learning-48d05f1759ae> (20.05.2018).

² Chatbot is a composite of the words “chat” and “bot” (robot).

³ See Helmut Leopold, Thomas Bleier, and FlorianSkopik, eds., *Cyber Attack Information System. Erfahrungen und Erkenntnisse aus der IKT-Sicherheitsforschung* (Wiesbaden, 2015).

⁴ See *Can Virtual Experiences Replace Reality? The future role for humans in delivering customer experience*, ed.

Oracle and Coleman Parkes (Redwood Shores, 2016); https://www.oracle.com/webfolder/s/delivery_production/docs/FY16h1/doc35/CXresearchVirtualExperiences.pdf

(May 20, 2018).

⁵ See “Gartner Says 25 Percent of Customer Service Operations Will Use Virtual Customer Assistants by 2020”; <https://www.gartner.com/newsroom/id/3858564> (May 20, 2018).

⁶ See “An indepth look at Google’s first Tensor Processing Unit (TPU)”; <https://cloud.google.com/blog/big-data/2017/05/an-in-depth-look-at-googles-first-tensor-processing-unit-tpu> (May 20, 2018).

⁷ For 5.4 billion us dollars.

⁸ See Nicole Hemsoth, “Intel, Nervana Shed Light on Deep Learning Chip Architecture,” *The Next Platform*, January 11, 2018; <https://www.nextplatform.com/2018/01/11/intel-nervana-shed-light-deep-learning-chip->

An increasing number of AI strategies are being pursued in a wide variety of applications today since they have repeatedly yielded far better results than traditional attempts at a solution. The best-known practical implementations are camera-based moving object detection in autonomous driving (a field in which Google has a clear innovation edge), speech recognition devices such as Alexa (Amazon), Siri (Apple), Bixby (Samsung), and Assistant (Google), Google’s All Language Translator, the boom of numerous chatbots² for a new kind of client interface for hotlines, online sales, and supply chain management, as well as AI algorithms for the protection of ever more complicated IT systems against cyberattacks.³

Chatbots will play an important economic part in this context. An Oracle study assumes that eighty per cent of businesses will rely on chatbots for customer service by 2020.⁴ Gartner confirms this assumption by predicting that more than half of us businesses will invest more capital in chatbots than in apps per year by 2021.⁵

Powerful hardware as the basis for AI

The breakthroughs of the impressive performance of AI algorithms built on computer-based neural networks (NN) are essentially owed to the availability of powerful processors and low-priced storage capacities; typically, several million parameters have to be continually computed for learning processes within the shortest time span.

Particularly the gaming industry with its demand for extremely powerful processors suited for high-resolution dynamic graphics and images plays a decisive part in the current breakthrough of AI technologies. Special graphics cards, so-called GPUs (graphic processing units), with processors were developed for video game consoles; arithmetic operations are efficiently carried out in parallel by these extremely powerful GPUs. Thanks to the gaming industry, which has ensured a high number of produced and sold items, powerful computers have become available and can be economically employed in different applications.

The development of powerful processors continues with quite some drive. Huawei has entered the international competition with its Mate 10 Pro and in-house processor Kirin 970. Huawei calls its processor NPU (neural processing unit); its computing power and speed are supposed to beat those of all GPUs produced to date. According to Huawei, the Kirin 970 features 5.5 billion transistors on one square centimeter. Comparable to human brain cells, more transistors mean more brainpower, which is why certain algorithms can be run in parallel.

Google calls its special AI hardware Tensor Processing Unit (TPU)⁶, IBM has its neuromorphic chip TrueNorth, AMD acquired the GPU specialist ATI as early as in 2006⁷, Intel bought Nervana und Mobileye, two companies specialized in deep learning processes, in 2016⁸. This sector of high-end processors for different markets sees an uninterrupted emergence of startups like KnuEdge, Eyeriss, krtkl, Graphcore, BrainChip, TeraDeep, Wave Computing, Horizon Robotics, NEURAM3, P-Neuro, or Spinnaker.⁹

Formen der künstlichen Intelligenz

Um Maschinen »intelligentes Verhalten« beizubringen, wurde lange mit regelbasierten Ansätzen gearbeitet, die durch übliche Programmstrukturen implementiert werden. Hierbei markierte der Einsatz mathematischer Methoden der Statistik und Wahrscheinlichkeitsrechnung bereits einen ersten wichtigen Entwicklungsschritt. Erste kommerzielle Erfolge von »intelligenten Systemen« erzielte man mit regelbasierten Expertensystemen in den 1980er-Jahren durch Verwendung von Datenbanken, logischen Entscheidungsbäumen und heuristischen Regeln wie statistischen Auswertungen, Zufallsstichproben und Ausschlussverfahren.¹⁰ Ein entscheidender Nachteil dieser regelbasierten Systeme ist der Umstand, dass die Regelbasis mit größeren Problemstellungen sehr rasch anwächst und dadurch unverständlich und immer schwerer wartbar wird.

Durch die Verfügbarkeit leistungsfähiger Prozessoren für eine große Anzahl paralleler Verarbeitungsschritte bekommen NN-Verfahren aktuell eine neue Bedeutung. Bei NN wird die Funktionsweise des menschlichen Gehirns nachgebildet und »Wissen« durch Verbindungsstrukturen zwischen Neuronen mit verschiedenen Gewichtungen und unterschiedlichen Regeln der Aktivierung je nach Signalzuständen an deren Verknüpfungen dargestellt. Lernen erfolgt durch die Veränderung von Verbindungen, durch deren Gewichtung und Stimulierungsregeln für die einzelnen Neuronen. Information wird auf mehreren Ebenen in Millionen Verbindungen »gespeichert«. Verbindungen zwischen den Neuronen werden mit von Experten beurteilten großen Mengen von Lerndaten eintrainiert. Die Verbindungen verändern sich durch die zunehmende Erfahrung. Solche Systeme werden Deep-Learning-Netzwerke genannt. Ein Modell der realen Welt wird nicht mehr wie bisher bei herkömmlichen Programmen durch eine mathematische Formel oder eine analytische Beschreibung erstellt, sondern durch Lerndaten erfahren und damit eintrainiert. Dies entspricht durchaus dem Konzept menschlichen Lernens.

Die KI-Grundlagenforschung widmet sich bereits weiteren Ansätzen der Entscheidungsfindung durch mathematische Verfahren der Mustererkennung, die den Maschinen selbstständiges Lernen aus vorhandenem Datenmaterial erlauben. Bei diesen »support vector machine«-Ansätzen werden bekannte Daten in Gruppen klassifiziert, um neuen Datensätzen dann durch einfaches Zuordnen zu bestimmten Gruppen eine Bedeutung zu geben, ohne dass der Inhalt der Daten erkannt wird oder Berechnungen durchgeführt werden.

¹⁰ Ein besonderes Beispiel österreichischer Hightechentwicklung ist eines der ersten Expertensysteme zur sicheren Steuerung von Eisenbahnstellwerken aus den 1990er-Jahren, das im Alcatel-ELIN-Forschungszentrum in Wien entwickelt wurde und sich bis heute am Weltmarkt als Technologie halten konnte (siehe Karl-Heinz Leitner, *Von der Idee zum Markt. Die 50 besten Innovationen Österreichs*, Wien 2003, S. 268 ff.).

architecture/
(May 20, 2018). See
also “Mobileye.
Das mächtige Auge des
Autos,” *Handelsblatt*,
March 13, 2017; <http://www.handelsblatt.com/impresum/nutzungshinweise/blocker/?callback=%2Funternehmen%2Findustrie%2Fmobileye-das-maechtige-auge-des-autos%2F19508880.html> (May 20, 2018).

⁹ See Shan Tang, “A List of Chip/IP for Deep Learning,” *Medium*, August 11, 2017; <https://medium.com/@shan.tang.g/a-list-of-chip-ip-for-deep-learning-48d05f1759ae> (May 20, 2018).

¹⁰ A striking example of Austrian high-tech development is provided by one of the first expert systems for the secure control of railway switch towers from the 1990s. It was developed in the Alcatel-ELIN research center in Vienna and has been able to assert itself as a technology on the world market to this day; see Karl-Heinz Leitner, *Von der Idee zum Markt. Die 50 besten Innovationen Österreichs* (Vienna, 2003), 268 ff.

Forms of artificial intelligence

For a long time, rule-based approaches implemented with the usual program structures were used for endowing machines with the ability to “behave intelligently.” The use of methods of mathematics applied to statistics and probability already marked a first decisive step of development. Rule-based expert systems employing databases, logical decision trees, and heuristic rules as well as statistical interpretations, random samples, and methods of elimination brought the first economic successes in the 1980s.¹⁰ A significant drawback of such rule-based systems lies in the fact that the base of rules grows very rapidly with greater problems and thus becomes incomprehensible and increasingly harder to service.

The availability of powerful processors for a large number of parallel processing steps currently lends NN procedures new significance. NNS copy the human brain’s manner of operation and represent “knowledge” by means of link structures between neurons with different priorities and rules of activation depending on the status of signals and their connections. Learning happens through alterations of connections, their weighting, and rules for stimulating individual neurons. Information is “stored” on several levels in millions of connections. A great many learning data assessed by experts are used to train the connections between neurons. These connections change with the accumulation of experience. Such systems have become known as deep learning networks. The model of the real world is no longer developed with the help of a mathematical formula or analytical descriptions, as it was in traditional programs, but experienced and thus trained through learning data. This corresponds perfectly to the concept of human learning.

Basic research conducted in the field of AI already explores further approaches to decision-making through mathematical procedures of pattern recognition that enable machines to learn independently from given data material. These so-called support vector machines classify known data into groups to give meaning to new datasets by simply assigning them to specific groups without reading the contents of the data or carrying out calculations.

Maschinelles Lernen, Lerndaten und Ground Truth

In Systemen, die auf NN aufbauen, erfolgt maschinelles Lernen durch Berechnungen der Parameter der **NEURONALEN NETZE** aufgrund von Lerndaten. Um zuverlässige Ergebnisse für echte Anwendungen zu erhalten, werden große Mengen von Lerndaten benötigt, wobei die Bedeutung der Daten für die jeweilige Anwendung – die sogenannte Ground Truth – bekannt sein muss.

Hier liegt eines der grundlegenden Probleme für den erfolgreichen Einsatz von KI-Technologien: Einerseits braucht man große Datenmengen, um die Algorithmen zu trainieren – was auch den Google-Ansatz erklärt, seit Jahren in verschiedenen Anwendungsbereichen Daten zu sammeln (für die Sensorik zum autonomen Fahren, für Übersetzungen von Text in die verschiedensten Sprachen, Sprach- und Audiodatensätze für Spracherkennungssysteme). Andererseits benötigt man Expertenwissen, um die vorhandenen Datensätze für den jeweiligen Anwendungsfall zu bewerten, zu »annotieren«. Für diese sogenannte Feature Extraction bedarf es unbedingt eines spezifischen Anwendungswissens, um vorhandene Daten entsprechend bewerten zu können und festzulegen, welche Parameter für welche Anwendungen relevant sind. Mit der Bewertung einer vorhandenen Datenmenge mit spezifischen Parametern legen Experten die Ground Truth fest, und die KI-Systeme können mit diesen Daten trainiert werden.

Nach der Feature Extraction durch Experten mit hohem Domänenwissen und der Festlegung der Ground Truth in Lerndaten ist die Modellbildung von geeigneten Netzstrukturen der dritte Bereich der Entwicklung der KI-Technologie. Es können Mechanismen unterschiedlicher NNS wie etwa rückgekoppelter neuronaler Netze (recurrent neural networks/RNN) verwendet werden. Berechnungen für Stimulierungsfunktionen der Neuronen können nach verschiedenen mathematischen Formeln durchgeführt werden.

Festzuhalten ist, dass ein Großteil der KI-Entwicklung im einfachen Ausprobieren von verschiedenen NN-Strukturen besteht: Nachdem die vorhandenen Lerndaten eingespielt worden sind, wird beobachtet, welche Ergebnisse mit einem bestimmten NN-Ansatz erzielt werden. Wenn das Ergebnis zufriedenstellend ist, geht man davon aus, dass die Anwendung im echten Einsatz korrekt funktionieren wird.

In vielen Bereichen sind solche von Experten beurteilten Datensätze wie jener des MAWI-Archivs für Internet-Verkehr¹¹ durch die globale Forschungscommunity entstanden oder werden mittlerweile – wie die Audiodatensätze zur Geräuscherkennung von Google (mehr als zwei Millionen von Menschen annotierte Zehn-Sekunden-Audiofiles)¹² – von großen Datenindustrien für Forschung und Entwicklung der KI zur Verfügung gestellt. Für viele neue Anwendungsgebiete ist es allerdings notwendig, Lerndaten zu entwickeln und durch Experten in aufwendiger Arbeit bewerten zu lassen.

NEURONALE NETZWERKE In künstlichen neuronalen Netzen werden in Computern bestimmte Eigenschaften von natürlichen Nervenzellen nachgeahmt – nämlich deren massive Verknüpfung untereinander, wobei die Stärke der Verknüpfung variabel ist. Lernen geschieht in diesen Netzen, die in mehreren Ebenen angeordnet sein können (Deep Learning), vor allem durch die Modifikation der Gewichte der Neuronen. ✕

¹¹ <https://en.wikipedia.org/wiki/MAWI> (20.05.2018).

¹² »A large-scale dataset of manually annotated audio events«; <https://research.google.com/audioset/> (20.05.2018).

Machine learning, learning data, and ground truth

Machine learning in NN-based systems is effected by calculating the parameters of **NEURAL NETWORKS** based on learning data. Large sets of learning data are required to achieve reliable results for genuine applications on condition that the meaning of the data for the respective application—the ground truth—must be known. This is one of the fundamental problems of the successful use of AI technologies. On the one hand, large amounts of data are necessary to train the algorithms—which also explains Google’s approach to collect data in different areas of application over many years (for its sensory technology in autonomous cars, for translations of text into various languages, as well as language and audio data-sets for speech recognition systems). On the other hand, expertise is needed to assess—to “annotate”—the given datasets for the application in question. This so-called feature extraction absolutely depends on specific application knowledge that allows assessing the data compatibly and defining which parameters are relevant for which applications. Evaluating a given amount of data with specific parameters, experts define the ground truth so that AI systems can be trained with these data.

Next to feature extraction carried out by experts with comprehensive domain knowledge and the definition of the ground truth in learning data, the construction of models of suitable network structures is the third field in the development of AI technologies. Mechanisms of different kinds of neural networks like recurrent neural networks (RNN) can be used. Calculations for the stimulation functions of neurons may draw on various mathematical formulae.

It should be noted that a large part of AI development consists in simply putting different NN structures to the test; after loading the given learning data, observation centers on which results are achieved with a certain NN approach. If the results are satisfactory, it can be assumed that the application will operate correctly in real mode.

Such evaluated datasets like those of the MAWI archive for Internet traffic¹¹ have been assembled by the global research community or are made accessible by big data businesses for research and development in the field of AI; Google’s audio datasets for sound identification (more than

two million ten-second audio files annotated by people) provide an impressive example.¹² Many new areas of application, however, require the development of learning data that have to be laboriously evaluated by experts.

¹¹ <https://en.wikipedia.org/wiki/MAWI> (May 20, 2018).

¹² “A large-scale dataset of manually annotated audio events”; <https://research.google.com/audioset/> (May 20, 2018).

NEURAL NETWORKS In artificial neural networks, specific characteristics of natural nerve cells are emulated in computers—namely, their massive interconnectedness, with the strength of the connection being variable. Learning occurs in these networks, which may be organized in several layers (deep learning), mainly by modifying neuron weights. ✕

Unsupervised Learning

Ein großes Potenzial im Bereich Deep Learning liegt bei den sogenannten unüberwachten («unsupervised») KI-Systemen, die mit der Fähigkeit ausgestattet sind, Objekte, Sprache und Bilder zu identifizieren, ohne zuvor mit einem verwandten Datensatz trainiert worden zu sein. Aus Sicht der Kontrolle von KI-Systemen durch den Menschen, also der Wahrnehmung der Letztverantwortung über Einsatzzweck und Zielverfolgung, besteht hier der größte Handlungsbedarf, damit die totale Autonomisierung der Maschinen sich nicht in nicht mehr verstehbarer und unlenkbarer Weise verselbständigt. In Zukunft werden in solchen Anwendungsszenarien wahrscheinlich »vortrainierte« Datensätze zum Einsatz kommen, mit denen das richtige Setup aufgesetzt und die Ergebnisse über Datenproben feingetunt werden. Dann allerdings könnten diese Ansätze neue Geschäftsmodelle im SaaS-Bereich (Software as a Service) hervorbringen, mit denen sich spezifische Kundenbedürfnisse im Datenmanagement exakt abbilden und Lernprozesse im laufenden Betrieb iterativ verbessern lassen. ✖

ÜBERWACHTES VS. UNÜBERWACHTES LERNEN

Beim überwachten Lernen bekommt ein neuronales Netz eine Rückmeldung durch den Menschen, ob es Strukturen in den Trainingsdaten richtig erkannt hat. Beim unüberwachten Lernen wird das System hingegen mit nicht-gekennzeichneten Daten gefüttert und muss selbstständig Strukturen erkennen und die Objekte z. B. in Cluster gruppieren. ✖

Unsupervised Learning

Great potential for the development of deep learning lies in so-called unsupervised AI systems. These systems are capable of identifying objects, language, and images without having been trained with a related dataset. Considering the control of AI, of the final responsibility for purpose of application and goal pursuit, such systems call for more action than others in order to prevent that the total automation of machines does not develop a momentum of its own which makes its operation incomprehensible and nondirigible. In the future, such application scenarios will probably work with “pretrained” datasets that allow establishing the right setup and finetuning the results through data samples. These strategies might lead to

new business models in the software-as-a-service sector (SaaS) that will make it possible to precisely render specific client demands in the data management and to iteratively improve learning processes during operation. ✖

SUPERVISED VS. UNSUPERVISED LEARNING In supervised learning, a neural network receives human feedback on whether structures were correctly identified in the training data. In unsupervised learning, the system is fed with non-labeled data and has to recognize structures on its own and, for example, group objects into clusters. ✖



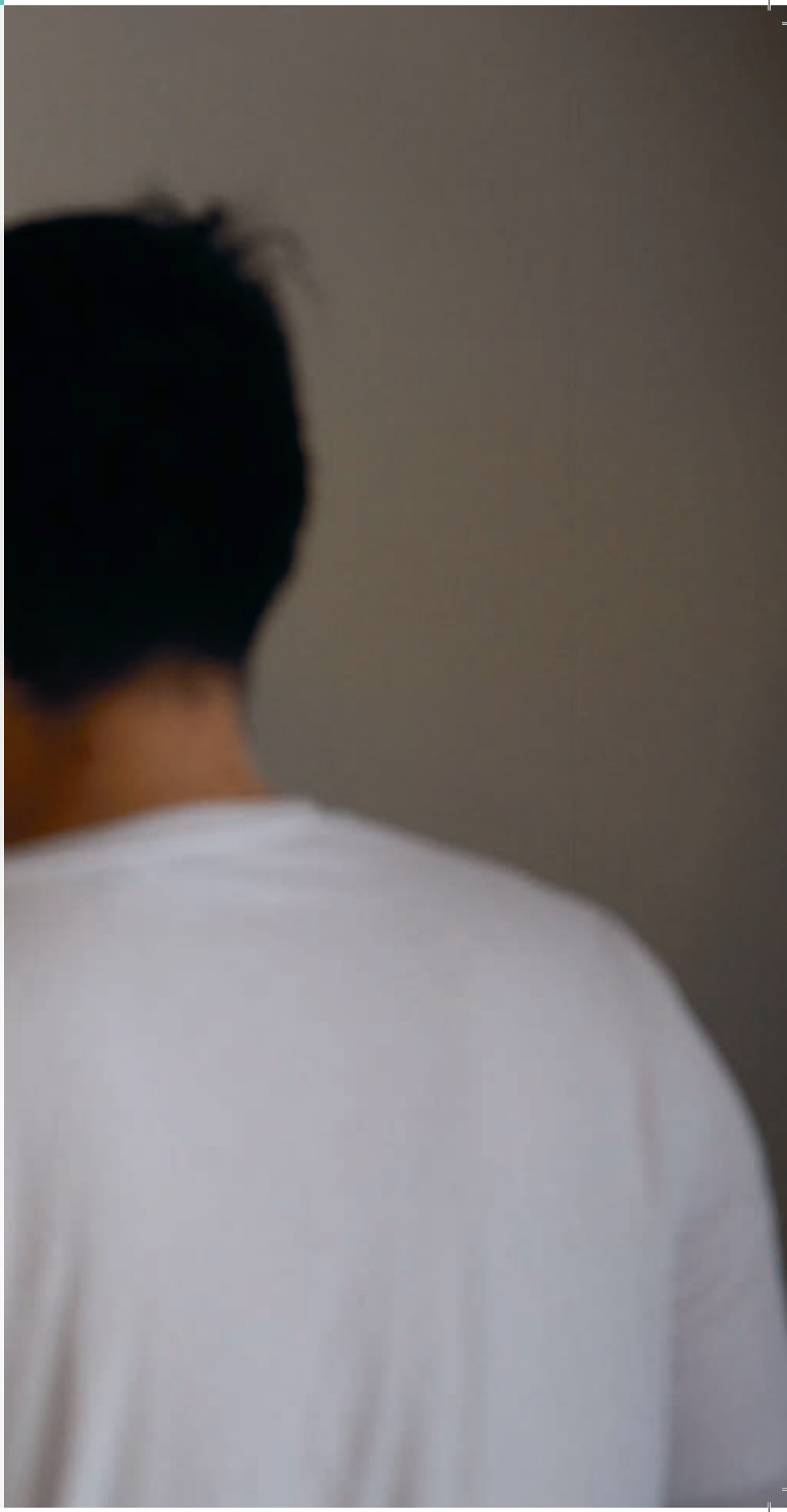
h
s
D
»l
d
v
l
S
B
it
a
it
B
G
D
a
u
s
e
“P
ir
v
fi
cl
W
it
a
a
a
tl
b
o
cl
le
to
—
©
F

hananona
STAIR Lab. (JP)

Die interaktive Arbeit »hananona« visualisiert die Kategorisierung von Blumen durch künstliche Intelligenz (KI). Sieht der Computer eine Blume, identifiziert er ihren Namen und zeigt auf einer »Blumenkarte« ihre Klasse an – was ein Bild vom Inneren des Gehirns der KI vermittelt. Das System erkennt artverwandte Exemplare und reiht sie in gemeinsame Bildergruppen ein.

“hananona” is an interactive work that visualizes how artificial intelligence (AI) classifies a flower. When it sees a flower, it identifies its name and shows its class on a visual “flower map”—a visualization of the inside of the AI brain. Each of the group of image clusters is a cluster of flower photos learned as belonging to the same class.

© Ars Electronica/
Florian Voggeneder



Manfred Tscheligi im Gespräch mit Martin Kugler

»Künstliche Intelligenz hat noch keinen Bauch«



Bei der Gestaltung von Systemen mit künstlicher Intelligenz wird derzeit zu wenig auf die Schnittstelle mit dem Menschen geachtet, meint Manfred Tscheligi. Das berge große Gefahren in sich, die von Unterforderung und Deskilling bis hin zu fehlender Akzeptanz der Technologie reichen. Durch überlegtes Design könnten viele dieser Probleme gemildert werden, ist der Spezialist für Mensch-Maschine-Interaktion überzeugt.

Wie fühlen sich Menschen, wenn sie einer intelligenten Maschine gegenüber treten?

Manfred Tscheligi: Um diese Frage zu beantworten, muss zuerst geklärt werden, was Intelligenz überhaupt ist. Einfach gesagt, handelt es sich bei einem System dann um ein intelligentes, wenn es Fähigkeiten hat, die nur der Mensch besitzt. Für jeden von uns ist Intelligenz etwas anderes: Es gibt beispielsweise, um nur einige Formen zu nennen, eine Problemlösungsintelligenz, eine soziale Intelligenz, eine kreative Intelligenz. Für intelligente Maschinen geht es darum, aus Daten, Wahrnehmungen und Stimuli etwas zu schaffen, was nicht schon völlig vorgegeben ist: etwa ein Problem zu lösen oder spezielle Kreativität einzubringen und das Ergebnis anschließend dem Menschen zu kommunizieren. Wenn nun der Mensch als intelligentes Wesen mit einem technischen System interagiert, stellt sich aus seiner Sicht folgende Frage: Verfügt das technische System über Intelligenz oder nicht? Merke ich überhaupt, dass es intelligent ist? Wie macht sich diese Intelligenz bemerkbar? Eines der großen Probleme in der Gestaltung solcher Systeme ist, dass deren Intelligenz nicht für den Menschen designt, sondern rein technisch realisiert wird. Es gibt keine Schnittstelle zwischen Maschinenintelligenz und Außen. Das System versucht gar nicht, seinem menschlichen Gegenüber beizubringen und verständlich zu machen, wie intelligent es ist. Dadurch kann es passieren, dass der Mensch gar nicht bemerkt, dass es sich um ein intelligentes System handelt bzw. wie dessen Intelligenz ausgeprägt ist.

Professor für Human-Computer Interaction an der Universität Salzburg, leitet das dortige Center for Human-Computer Interaction sowie das Center for Technology Experience am AIT. Studierte Wirtschaftsinformatik; Doktorat der Sozial- und Wirtschaftswissenschaften. Mittlerweile ist er seit mehr als 20 Jahren in den Bereichen interaktive Systeme, Human-Computer Interaction, Usability Engineering, User Interface Design und User Experience Research tätig und gilt als Pionier dieses Gebietes in Österreich innerhalb der Ausbildung, als Forschungsgebiet sowie in der industriellen Anwendung.

Manfred Tscheligi, Professor for Human-Computer Interaction at the University of Salzburg, Head of its Center for Human-Computer Interaction, and Head of the Center for Technology Experience at the AIT. Manfred Tscheligi studied business informatics and holds a doctorate in social science and economics. He has been active in the areas of interactive systems, human-computer interaction, usability engineering, user interface design, and user experience research for more than twenty years. He is considered a pioneer in this area in Austria, with regard to not only education, but also research and industrial application.

Manfred Tscheligi in an interview with Martin Kugler

“Artificial intelligence still lacks gut instincts”

When it comes to designing artificial intelligence systems, the computer-human interface is still being neglected, says Manfred Tscheligi. According to him, this carries the danger of underchallenge and “deskilling” and may therefore prevent the technology from becoming accepted. A well-thought-out design could help mitigate many of these problems, the expert on human-computer interaction is convinced.

How do humans respond when confronted with an intelligent machine?

Manfred Tscheligi: In order to be able to answer this question, it must be clarified at first what intelligence actually is. To put it simply, a system is intelligent if it has skills that only humans have. For each of us, intelligence is something different: for example, to name just a few of its manifestations, we know problem-solving intelligence, social intelligence, creative intelligence. Intelligent machines are expected to create something not yet completely predefined from data, perception, or stimuli, such as solving a problem or contributing with a particular form of creativity and subsequently communicating the results to humans. From their perspective, when interacting with a technological system, humans as intelligent beings will ask themselves: Is this technological system intelligent or not? Will I even recognize whether it is intelligent? How does its intelligence show? One of the fundamental problems in the development of such systems is that their intelligence is not designed for humans but is realized purely technically. There’s no interface between machine intelligence and the outside world. The system doesn’t even try to explain to its human counterpart how intelligent it is. It may therefore happen that people won’t even notice that they are interacting with an intelligent system or in what ways its intelligence is distinctive.

Warum sehen Sie diese rein technische Realisierung als Problem an?

MT: Als Mensch habe ich das Grundbedürfnis, das technische System zu beherrschen – und nicht anders herum. Eine Maschine soll mich nur bei einer bestimmten Aufgabe unterstützen und mir nicht ebenbürtig sein. Bei der Mensch-Maschine-Kommunikation ist es wichtig, dass das System so gestaltet ist, dass der Nutzer das System versteht und die volle Kontrolle darüber hat und behält. Es muss klar sein, was das System zulässt und wie intelligent es wirklich ist. Dieser Aspekt wird derzeit zu wenig berücksichtigt. Dabei handelt es sich um eine Designfrage. Momentan ist die Entwicklung von künstlich intelligenten Systemen sehr von der Technik bestimmt und vernachlässigt die Gestaltung noch zu sehr.

Können Sie das an einem Beispiel erläutern?

MT: Ein gutes Beispiel sind autonome Fahrzeuge. Das sind de facto Roboter, und zwar technisch gesehen oft schon bessere Roboter als Industrieroboter. Man sitzt also genau genommen in einem Roboter. Ein solches Auto hat offensichtlich gewisse intelligente Verhaltensweisen: Die künstliche Intelligenz sorgt dafür, dass das Auto allein fahren kann und auf Verkehr und Umgebung richtig reagiert. Die Frage ist: Verlasse ich mich auf die Technologie? Will ich ihr ausgeliefert sein? Und weiß ich überhaupt, was das System wirklich kann? Derzeit gibt es keine Möglichkeit, das System zu fragen. Im Vergleich dazu kann ich einem Menschen jederzeit die Frage stellen, ob er dies oder jenes kann. Gibt man einem Menschen beispielsweise ein Rätsel auf, merkt man relativ schnell, in welcher Verfassung er sich befindet und ob er imstande ist, das Rätsel zu lösen. Gibt man einer Maschine eine Aufgabe zu lösen, erkennt man nicht so leicht, ob das System dabei ist, das Rätsel zu lösen, oder ob es nicht in der Lage dazu ist.

Viele Menschen testen ja auch absichtlich aus, was eine Maschine kann...

MT: Ja, dabei handelt es sich um eine Art Machtkampf. Tut das System vielleicht nur so, als wäre es intelligent? Der Mensch wird dann erst recht Fehler suchen, herausfinden wollen, was die Maschine kann und woran sie scheitert. Schließlich stellt sich auch die Frage, wie viel künstliche Intelligenz ein Individuum verträgt. Als Mensch will man oft einfach keinen Superstar, der alles kann, als Partner – das ist zutiefst in uns verankert. Laut Studien akzeptiert man technische Systeme eher, wenn sie einem sagen, was sie können, aber auch wozu sie nicht in der Lage sind. Das ist ein Paradoxon: Ich akzeptiere ein System eher, wenn es nicht perfekt ist – obwohl es dadurch ja weniger kann.

Why do you consider this purely technical realization a problem?

MT: Being human, I basically feel the need to be in control of a technological system instead of the other way round. A machine is supposed to support me in a specific task, it's not meant to be my equal. When it comes to human-computer communication, it's crucial that the system has been designed in such a way that users will understand it and be able to master it and keep it under control. It must be clear what the system will tolerate and how intelligent it really is. This aspect is currently not taken sufficiently into account, although it's purely a matter of design. Presently, the development of intelligent systems is overdetermined from a technological point of view while the design is still being neglected.

Can you give us an example?

MT: Autonomous vehicles are a perfect example. They are in fact robots and in terms of technology frequently even better than industrial robots already. So you are actually driven around by a robot. This type of car evidently exhibits certain forms of intelligent behavior: artificial intelligence ensures that the car can drive autonomously and respond properly to traffic and to its surroundings. What is at issue is whether I'm willing to rely on this technology. Do I wish to be at its mercy? And do I know what the system is actually capable of? It's currently not possible to ask the system. On the other hand, I can always ask humans if they are capable of accomplishing certain tasks. For instance, when confronting a person with a riddle you'll notice relatively soon in what state of mind they are and if they are in a position to actually solve it. When asking a machine to solve a problem, it's not that easy to recognize whether the system is in the process of solving it or if it is unable to do so.

Many people put a machine to the test deliberately to find out what it can do...

MT: Yes, this boils down to a kind of power struggle. Maybe the system merely pretends to be intelligent? Humans will then make a point of trying to find out what a machine can do and where it fails. Finally, there is also the question of how much artificial intelligence a human can cope with. As a human, you frequently simply don't want to have a superstar capable of doing everything as your partner—this is something that is deeply rooted in man. According to certain studies, it is

Wie müsste die Schnittstelle also gestaltet sein, damit Menschen ein System akzeptieren?

MT: Die Systeme müssen sagen, wenn sie etwas nicht können. Nehmen wir als Beispiel ein Navigationssystem im Auto: Es tut so, als könne und wisse es alles. Kein System sagt mir: Ich kenne mich da nicht aus. Das System weiß zwar, dass es zu wenig Daten hat, um verlässliche Aussagen zu treffen, vermittelt das jedoch nicht. Stattdessen bringt mich das Navi irgendwohin – zum Beispiel in die Nebenfahrbahn oder auf den Acker. Hier geht es also um eine Form von sozialer Intelligenz. Dazu gehört bei uns Menschen, dass wir manchmal auch sagen: Ich habe keine Ahnung, ich weiß es nicht, da muss ich nachschauen, etc. Dieses Eingestehen von Nichtkönnen ist etwa bei autonomen Fahrzeugen ein entscheidender Aspekt, denn das Fahrzeug suggeriert mir, dass es jede Situation meistern kann. Was geschieht aber, wenn ein Sensor ausfällt? Sagt mir das Auto dann, dass es zu wenige Daten hat? Stellen wir uns zwei Szenarien vor. Vor Antritt der Fahrt sagt mir das Auto: Liebe Autonome-Auto-Fahrerin, lieber Autonome-Auto-Fahrer, bitte pass auf, es könnte sein, dass während unserer Fahrt aufgrund des dichten Schneefalls ein Sensor ausfällt! Mit dieser Information fährt man ganz anders Auto als im zweiten Szenario, wenn einem das Auto von Anfang an suggeriert: Fahr einfach los! Dann verlasse ich mich auf die Maschine und gehe davon aus, dass alles funktioniert. Kommt es jedoch zu einem Ausfall des Systems und der Mensch muss plötzlich das Steuer übernehmen, kann es riskant werden, da er als »Beifahrer« des autonomen Fahrzeugs vorerst nicht ganz aufmerksam ist. Dieses »Hand-Over«-Problem ist eine extrem heikle Sache. Deswegen ist es wichtig, wie sich eine Intelligenz – oder oft »Pseudointelligenz« – dem Menschen gegenüber präsentiert. Wenn das System von Menschen ernst genommen werden soll, muss es ehrlich sein, muss es in einem gewissen Sinn menschlich sein. Wenn man Menschen gegenübertritt, will man, dass das Gegenüber ehrlich ist. Diese Kommunikationsintelligenz spielt auch beim Umgang von Menschen mit Maschinen eine wesentliche Rolle.

Verstehe ich das richtig: Man sollte eine Maschine also so designen, dass sie dem Menschen zeigt, dass sie nicht perfekt ist?

MT: Ja – auch um dem Ganzen mehr Seele, mehr Emotion zu geben. Ich würde es zum Beispiel gut finden, wenn das autonome Auto, nachdem man eingestiegen ist, sagt: »Heute fahren wir nicht, denn heute ist es zu gefährlich. Fahr lieber mit der Straßenbahn!« Das wäre intelligent! Sehr eng damit hängt ein weiteres Thema zusammen: Bei autonomen Systemen wird es wichtiger werden, zu über-

easier for us to accept technological systems if they can tell us both what they can do and what they cannot do. This is a paradox: we rather tend to accept a system when it's imperfect—even if it can do less then.

So how should the interface be designed for humans to accept a system?

MT: The system should be able to tell me what it cannot do. Take the satnav systems used in cars, for example. They behave as if they were capable of doing and knowing everything. There is no system telling me: I'm at a loss here. The system knows very well in fact that the data at its disposal is insufficient to make a reliable statement, but this is not what it communicates. Instead, the satnav takes me to some strange place, like a frontage street or a field. What would be needed is a form of social intelligence, which for us humans involves admitting now and then that we have no idea or must first check, etc. This confession of inability is a crucial aspect about autonomous vehicles, for instance, as the vehicle suggests that it is capable of mastering any situation. What will happen if one of its sensors becomes inoperative? Will the car let me know then that the data available is inadequate? Let's imagine two different scenarios. Before setting off, the car will tell me: Dear autonomous car driver, please note that due to heavy snowfall one of the sensors might malfunction during our ride! Knowing this, the way you'll be driving the car will be different from the way you'll be driving it if the car suggests from the very beginning: just start! Then I'll rely on the machine and assume that everything will be working fine. However, if the system does fail then and the human driver is forced to take over, the situation can become risky because as the "co-driver" of an autonomous car a person will not be paying that much attention at first. This hand-over problem is a matter of extreme delicacy. It is therefore vital how an intelligent—or frequently "pseudo-intelligent"—system presents itself to its human counterpart. If the system wants humans to take it seriously, it must be honest; in a certain sense, it must be human, too. When facing another person, you expect them to be honest. This communicative intelligence also plays a major role in the interaction between man and machine.

prüfen, wie es dem Nutzer geht, wie es um seine persönliche Verfassung bestellt ist. Stellen Sie sich wieder das Szenario von vorhin vor: Sie sitzen in einem autonomen Fahrzeug und müssen plötzlich das Lenkrad übernehmen, da ein Sensor ausgefallen ist; allerdings waren sie davor auf einer Feier und fühlen sich eigentlich nicht imstande, das Fahrzeug zu lenken. Das bedeutet, dass man sich nur dann in ein autonomes Auto setzen darf, wenn man vorher einen Alkoholtest gemacht hat. Man muss dem System eine Intelligenz geben, die feststellen kann, wer es unter welchen Umständen und in welchem Zustand benutzen darf. Das gibt es zwar jetzt auch schon – nur entscheidet das jetzt der Mensch. Wenn man dem System die Verantwortung gibt, muss man ihm auch diese Möglichkeit geben. Darfst du überhaupt übernehmen? Je autonomer und intelligenter Systeme werden, umso wichtiger wird diese Verwendungsentelligenz aus der Mensch-Maschine-Sichtweise.

Es heißt immer wieder, dass uns die Technik dümmer macht, wenn sie uns mehr und mehr Aufgaben abnimmt.

MT: Wenn eine Maschine alles automatisch macht oder wenn ein Auto von selbst fährt, nimmt die Aufmerksamkeit ab, muss man doch selbst nichts aktiv tun. Wenn viele intelligente Systeme um uns herum sind, muss man nicht mehr selbst denken. Dadurch verlernt man auch viele Dinge, der Mensch entwickelt sich zurück. Wer kann heute zum Beispiel noch kopfrechnen, da jedes Handy einen Taschenrechner eingebaut hat? Selbst wenn man es einmal gelernt hat, ist man nicht mehr darin geübt. Das nennt man in der Fachsprache Deskillung. Wenn man nicht mehr gewohnt ist, etwas selber zu tun, kann man nicht mehr wissen, ob das Ergebnis richtig ist. Fehler werden nicht mehr erkannt, und man ist den Systemen ausgeliefert. Dies birgt ebenfalls Gefahrenpotenzial, weil man nicht mehr ohne Systeme auskommt und auf unvorhergesehene Geschehnisse nicht entsprechend reagieren kann.

Kann man dagegen etwas tun?

MT: Intelligente Systeme sollten den Nutzern hin und wieder neue Herausforderungen bieten und sie (über-)fordern. Dabei handelt es sich wiederum um eine Designfrage, die direkt in die Systeme integriert werden muss. Aktuell gibt es keine Tendenz in diese Richtung. Man könnte Systeme beispielsweise so designen, dass sie uns Aufgaben stellen, die uns zum Denken bewegen – das geschieht auch in der Kommunikation zwischen Menschen. Man könnte aber in technischen Systemen auch bewusst Fehler anlegen, um den Menschen zum Mitdenken anzuregen und ihn dadurch

If I understand what you said correctly, one should thus design a machine that demonstrates its imperfection to humans?

MT: Yes—also to give it more soul, to make the procedure more emotional. For example, I'd appreciate it if, once I've climbed into my autonomous car, it'd tell me: "We won't be driving together today because it's too dangerous. You'd better take the streetcar!" This would indeed be intelligent! This also takes us to a closely related subject: it'll become increasingly important for autonomous systems to see how the user is doing and check up on his or her personal condition. Imagine the scenario described above: You sit in an autonomous car and suddenly must take the steering wheel because a sensor has become inoperative; but having been to a party, you actually don't feel up to driving. This means that you should only be allowed to drive an autonomous car if you've taken a breath test before. The system therefore has to be furnished with an intelligence capable of determining who is authorized to use it and under what circumstances and in what condition. This does exist already—but the decisions are still made by humans. If responsibility is delegated to the system, it must also be given the chance to decide. Are you in a position to take over? The more autonomous and intelligent the system, the more important this operational intelligence will become from the human-machine perspective.

It's often said that technology is making us dumber as it's taking more and more tasks off our shoulders.

MT: If a machine does the job automatically, or if a car drives autonomously, our attention will decrease since we don't need to contribute actively. Surrounded by a large number of intelligent systems, we are no longer forced to think ourselves. This is why we tend to forget many things—humankind is regressing. Who is still able to do mental math now that every cell phone comes with an integrated pocket calculator? Even if you learned it once, you'll now lack the practical experience. In expert language, this is called "deskilling." If you're not used to doing something yourself, you'll no longer know if the result is correct. You'll fail to identify errors and will be completely dependent on systems. This harbors potential risks, as you can no longer do without these systems and cannot respond adequately to unexpected events.

zu trainieren. Ein autonomes Auto könnte zum Beispiel einmal sagen: »Heute mag ich nicht!« Das wäre eine ganz neue Situation für den betroffenen Menschen und würde ihn zum Nachdenken anregen. Im ersten Moment würde man sich eventuell ärgern. Aber dann wird man sich zu überlegen beginnen, warum das System das macht, und versuchen, eine Lösung zu finden. Wichtig dabei ist natürlich, dass das System eine Begründung gibt. Das wäre auch eine Art von Intelligenz.

Was passiert, wenn Menschen sich einer übermächtigen Intelligenz ausgesetzt fühlen?

MT: Neben dem Deskillung-Problem gibt es das bereits erwähnte Phänomen des Kontrollverlusts. Das ist psychologisch sehr wichtig. Daher haben zum Beispiel Webseiten eine »Home«-Taste: Sobald man sie drückt, kommt man »nach Hause« und findet sich in einer gewohnten Umgebung wieder. Kontrollverlust führt zu Panik, zu Unsicherheit und zu Ineffizienz, weil man plötzlich über Dinge nachdenkt, die normalerweise automatisch ablaufen. Wenn der Mensch die Kontrolle über den »next step« verliert, wird es kritisch. Wenn man nicht mehr weiß, was geschieht, drückt man lieber den Aus-Knopf. Kontrollverlust führt daher auch zu »Non-Use«. Wer Angst hat, dass ein Roboter seinen Job gefährdet, ist versucht, den Roboter als nicht funktionierend darzustellen, oder geht sogar einen Schritt weiter und sabotiert die Maschine. Auch das ist eine Art »Non-Use« von Technologie.

Das geht natürlich völlig am Zweck der Technologie vorbei.

MT: Das Ziel muss sein, solche Situationen zu verhindern. Beim Design von Systemen und bei der Form der Kommunikation muss stärker in Betracht gezogen werden, wie sich Systeme dem Menschen gegenüber verhalten. Ein Schlagwort ist hier »honest design«: eine ehrliche Darstellung, wie man mit dem System umgehen kann und soll. Da sind wir wieder bei der sozialen Intelligenz: Man muss wissen, woran man ist. Man soll Hilfe bekommen, wenn man sie braucht, man muss sich aber auch zurückziehen können, wenn man sie nicht will. Diese soziale Intelligenz muss in Systeme integriert werden, das passiert nicht automatisch.

Can something be done against it?

MT: Now and then, intelligent systems should confront users with new challenges and overtax them. This, too, is a matter of design and has to be directly integrated into the system. Currently, however, there's no such tendency in this direction. For example, systems could be designed in such a way that they confront us with problems and get us thinking—which also normally happens in interhuman communication. Or you could also deliberately build mistakes into technological systems so as to stimulate people's mind and keep them in training. An autonomous car might say: "I don't feel like driving today!" This would confront users with an entirely new situation and demand their initiative. At first, you'll probably get angry. Then you'll start wondering why the system has reacted like that and try to find a solution. Of course, it'll be essential that the system should give you some reason for its response. This would also be a form of intelligence.

What happens when people feel exposed to some all-powerful intelligence?

MT: Besides the problem of "deskilling," there is also the above-mentioned phenomenon of a loss of control. This is very important from a psychological point of view. That is why websites have a "home" button, for example: as soon as you click it, you'll be "taken home" and return to familiar surroundings. A loss of control leads to panic, uncertainty, and inefficiency, because you'll suddenly start thinking about things that usually happen automatically. If a person loses control of the "next step," this could end up in a critical situation. If you are in the dark about what's going to happen next, you'd better hit the off button. A loss of control will also entail "non-use." Those afraid that a robot might take their jobs will try to present the robot as malfunctioning or even go further and sabotage the machine. This is also a form of "non-use" of technology.

This completely defeats the purpose of technology, of course.

MT: The goal must be to prevent such situations. When designing systems and deciding on the form of communication, we must increasingly consider how systems interact with humans. "Honest design" is an appropriate catchword in this context: an honest representation of how one can and should deal with the system. This takes us back to social intelligence: one must know what to expect. We should obtain help if it is needed, but

Wie weit ist die Forschung beim Design von künstlich intelligenten Systemen?

MT: Es gibt viele Studien über diese Probleme, und es wird immer klarer, für welche Phänomene man Lösungen benötigt. Antworten haben wir allerdings noch nicht viele. Welche Designprinzipien gibt es dafür? Wie setzt man diese um? Was wirkt? Was ist nachhaltig? Weitere wichtige Fragen sind: Wie wirken diese Designprinzipien auf bestimmte Gruppen? Und wie wirken sie auf das Individuum – im Sinn einer Personifizierung? Es geht um die Bedürfnisse, die ein Mensch hat, um seinen individuellen Charakter. Im Ergebnis muss das dahin führen, dass jeder seine eigenen Designs bekommt. Das ist kommerziell aufwendig, und dazu benötigt man auch intelligente Systeme, die das Design intelligenter Systeme auf die individuellen Bedürfnisse intelligenter Menschen ausrichten. Das versucht man bereits im Consumer-Bereich: Einkaufssysteme wissen, was Menschen interessiert. Aber für das Design von Systemen und für deren Kommunikation mit Menschen werden solche Methoden noch kaum eingesetzt. Die Daten dafür wären da: Jeder Computer hat eine Kamera, man könnte etwa in jedes System eine Eye-Tracking-Vorrichtung einbauen und Analysen durchführen, welche Stelle eines Bildschirms ein Mensch immer im Auge hat: Dort könnte man die für ihn wichtigen Bedienfelder platzieren – und dafür andere Elemente, die er nie benutzt und die ihn nur stören, aus seinem Blickfeld entfernen. Oder man könnte für eine sehgeschwachen Person, die eine Brille aufsetzt, automatisch die Schrift größer machen. So könnte man die Systeme viel stärker personalisieren.

Ist aus Ihrer Sicht die Sorge vieler Menschen berechtigt, dass Maschinen irgendwann intelligenter sein werden als Menschen?

MT: Es wird immer weniger Aspekte und Situationen geben, in denen der Mensch besser ist. Maschinen haben durch die Datenmengen, die sie verarbeiten könnten, teilweise enorme Vorteile bei Aufgaben, bei denen es um Wissen, um ein Abgreifen von Datenbanken, um Verknüpfungen etc. geht. Maschinen werden in diesen Bereichen immer besser, der Mensch wird ein Problem haben, dabei mitzukommen. Auch autonome Autos werden immer besser werden. Aber sie werden letzten Endes dennoch nicht so intelligent sein wie Menschen. Es wird immer Situationen geben, in denen der Mensch besser ist – etwa Situationen, auf die man nicht vorbereitet ist, oder unklare Situationen.

one should also be able to withdraw if it is unwanted. This type of social intelligence must be integrated into systems, it doesn't happen automatically.

How far has research progressed into the design of artificial intelligence systems?

MT: There are numerous studies dealing with these problems, and it's becoming increasingly clearer what phenomena require solutions. However, there are not yet that many answers. What are the design principles? How can they be implemented? What will work? What about sustainability? Further important questions are: How do these design principles impact on specific groups? And how do they impact on the individual—in the sense of personification? What matters are human needs, people's individual characters. This must lead to everyone receiving their own personal design in the end. This is demanding commercially and requires intelligent systems capable of adapting the design of intelligent systems to the individual needs of intelligent people. Attempts have already been made in the consumer field: shopping systems know what people are interested in. However, these methods are hardly employed yet for the design of systems and their communication with humans. The necessary data would in fact be available; every computer has a camera installed so that each system could be equipped with an eye tracking device to analyze which areas of the screen are constantly in the user's focus: control panels could be placed in these areas while other elements that are never used or just irritating could be removed from the user's focus. The type size could also be automatically enlarged for a visually impaired person putting on eyeglasses.

Are the worries many people have about machines becoming more intelligent than humans one day justified from your perspective?

MT: There'll be fewer and fewer aspects and situations in which humans will outstrip machines. Because of the huge amount of data available to machines, they are partly at an enormous advantage when it comes to retrieving and processing knowledge, reading databases, linking information, etc. As machines are constantly improving in these disciplines, it'll be difficult for humans to keep up with them. Autonomous cars will also improve. Yet they won't be as intelligent as humans are eventually. There'll always be situations in which humans are superior—such as unexpected or ambiguous situations.

Es wird also Dinge geben, die Menschen vorbehalten bleiben?

MT: Ja – etwa Erfahrungswerte oder ein gewisses Gespür. Wenn man Manager befragt, hört man oft, dass viele Entscheidungen aus einem Gefühl heraus getroffen werden, dass vieles eine Bauchentscheidung ist. Künstliche Intelligenz hat noch keinen Bauch. Gefühl ist nicht algorithmisch erfassbar, weshalb Maschinen hier immer einen Nachteil haben werden. Im Bereich der Kreativität wird das noch eine Zeit dauern – heutige Maschinen sind nicht kreativ. Ähnliches gilt für soziale Situationen: Wenn mehrere Menschen zusammensitzen und über Dinge reden, wenn ein Wort das andere ergibt, dann hat man plötzlich eine gemeinsame Idee, aus der etwas entsteht. Man stelle sich vor, dass sich drei Maschinen zusammensetzen: Sie werden (noch) keine gute Idee entwerfen können. Auch das ist Intelligenz. Da gewinnen die Menschen.

Sollte man Ihrer Meinung nach manche Dinge bewusst Menschen überlassen?

MT: Es muss Dinge geben, bei denen der Mensch nicht durch eine Maschine ersetzt wird. Ich möchte nicht in einer Welt leben, die nur aus Maschinen besteht. Im Gegenteil: Man sollte ausdrücklich feststellen, in welchen Situationen der Einsatz von Technologie sinnvoll ist – etwa bei Sicherheitssystemen, Routinetätigkeiten oder schweren Verrichtungen – und was dem Menschen überlassen wird. Wenn man Menschen mit Technologie überflutet und sagt, dass alles automatisiert wird, führt das zu Angst, was wiederum starke Akzeptanzprobleme mit sich bringt. Die gesellschaftliche Digitalisierungsdiskussion muss in dieser Hinsicht wesentlich differenzierter werden. Es bedarf einer Synergie zwischen der menschlichen Intelligenz und gut eingesetzten künstlich intelligenten Systemen – und zwar dort, wo man sie wirklich brauchen kann, wo sie uns helfen und wo wir selbst entscheiden können, ob man sie braucht oder nicht. Es geht um die Synergie Mensch und Maschine und nicht um ein Entweder-Oder! ✕

So there'll be domains that will continue to be reserved for humans?

MT: Yes—such as empirical knowledge or a certain sense of intuition. When interviewing managers one is often told that many decisions are made intuitively, from the gut. Artificial intelligence still lacks gut instincts. Emotions cannot be ascertained algorithmically, which is why machines will always fall behind in this respect. As far as creativity is concerned, it will still take a while—today's machines are not yet creative. This similarly goes for social settings, with several people meeting and talking things over, with one word leading to another until, all of a sudden, they'll arrive at an idea together from which to create something new. Imagine three machines conversing: they'll not (yet) be capable of conceiving an excellent idea. This also has to do with intelligence. In this regard, humans still emerge victorious.

Do you think that some things should deliberately be left up to humans?

MT: There absolutely must be things where humans are not replaced by machines. I don't want to live in a world consisting solely of machines. On the contrary, it should explicitly be defined in what situations the use of technology makes sense—such as for security systems, routine jobs, or heavy labor—and what tasks should be left for humans. If humans are overwhelmed with technology and told that everything will be automated, this will raise fears, which in turn will lead to serious problems of acceptance. Society's discussion of digitization must become much more differentiated in this respect. What is required is a synergy between human intelligence and cleverly employed artificial intelligence systems—in those areas where they are truly needed, where they support us, and where we can decide for ourselves if they are of use or not. This is about the synergy of humans and machines; it's not either/or but both! ✖

Zwischen Utopie und Dystopie

Mit Wörtern wie »technologische Singularität«, »Superintelligenz« oder »Transhumanismus« wurde eine Reihe von Begriffen geschaffen, welche die aktuellen Debatten über künstliche Intelligenz prägen. – Ein kurzer Rundgang durch alte und neue Befürchtungen.

Es war einmal: Angeblich hat der Prager Rabbiner Judah Löw am 17. März 1580 aus einem Klumpen Ton eine menschliche Figur geformt und ihr in einem kabbalistischen Ritual Leben eingehaucht. Dieser Golem sollte der jüdischen Gemeinde als Wächter gegen die Bedrohungen durch die christliche Mehrheitsbevölkerung Prags helfen. In der Stube des Rabbi pflegte der Golem in einer Ecke zu sitzen, heißt es in der Überlieferung. Zum Leben erweckt wurde er erst dadurch, dass ihm ein Zettel mit dem Namen Gottes unter die Zunge gelegt wurde. Die Entfernung dieseszettels entzog ihm seine Lebensenergie. Doch eines Tages vergaß der Rabbi, ihm den Zettel aus dem Mund zu nehmen – und da begann der Golem durch die Straßen des Prager Ghettos zu rasen und alles zu zerschlagen, was sich ihm in den Weg stellte. »Da warf sich der Rabbi vor ihn, entfernte den Zettel und vernichtete diesen, woraufhin der Golem in Stücke zerfiel«, heißt es in der Legende.

Eine andere Version beschreibt, wie die Frau des Rabbi Löw gegen dessen ausdrückliches Geheiß, dass der Golem für derartige Arbeiten nicht heranzuziehen sei, dem Golem befahl, Wasser ins Haus zu bringen. Dann ging sie auf den Markt, und der Golem trug weiter mehr und mehr Wasser ins Haus – weil ihm nicht befohlen worden war, damit aufzuhören. Möglicherweise diene diese Legende als Vorlage für Johann Wolfgang von Goethes Ballade vom »Zauberlehrling« (die Thematik ist freilich bis in die Antike zu Lukians »Die Lügenfreunde oder: Der Ungläubige« zurückverfolgbar). Jedenfalls: Der Besen, der einmal zum Leben erweckt worden war, wollte nicht mehr zu arbeiten aufhören – was in dem Vers »Die ich rief, die Geister/werd ich nun nicht los« zum geflügelten Wort wurde.

Between Utopia and Dystopia

Words like “technological singularity,” “superintelligence,” and “trans-humanism” have been invented within the context of a number of concepts dominating the current discourse about artificial intelligence. Let us take a look at old and new fears.

Once upon a time ... The story goes that on March 17, 1580 Rabbi Judah Loew of Prague created a human figure from a lump of clay and breathed life into it in a cabalist ritual. This golem should help protect the Jewish community as a guardian from threats coming from the Christian majority of Prague’s population. As tradition has it, the golem used to sit in a corner of the rabbi’s studio, only coming to life when a piece of paper inscribed with the name of God was placed under his tongue. Removing this piece of paper deprived him of his vitality. But one day the rabbi forgot to take the piece of paper out of his mouth—which resulted in the golem’s raging through the streets of the Prague ghetto and smashing everything in his way. According to legend, “the rabbi, throwing himself in front of him, removed the piece of paper and destroyed it, whereupon the golem came to pieces.”

Another version describes how Rabbi Loew’s wife, against his explicit instructions not to use the golem for such chores, ordered him to fetch water and bring it into the house. Then she went to the market whilst the golem was fetching more and more water as no one had told him to stop. This legend might have inspired Johann Wolfgang von Goethe’s ballad “The Sorcerer’s Apprentice” (although the theme can be traced back as far as antiquity, namely to Lucian’s “The Lover of Lies,” also known as “The Doubter”). In any case, the broomstick, once brought to life, refused to stop working—which became a familiar dictum through the verse “Spirits that I’ve cited/my commands ignore.”

Von der fantastischen Literatur...

Figuren wie der Golem oder der Zauberlehrling, die der Kontrolle des Menschen entgleiten, haben die Fantasie der Menschen schon immer angeregt und Dystopien heraufbeschworen – etwa in Mary Shelleys *Frankenstein* aus dem Jahr 1818 und vielen anderen Geschichten der fantastischen Literatur des 19. Jahrhunderts. Mit dem Aufstieg der modernen Wissenschaft wurde der technologische Einschlag dieser Geschichten immer stärker; zudem wurden die Geschichten vermehrt auch mit politischen Ideen verknüpft. Man denke etwa an Bücher wie Aldous Huxleys *Schöne neue Welt* (1932), George Orwells *1984* (1948) und Ray Bradburys *Fahrenheit 451* (1953) oder an die autokratische Maschinenwelt in Fritz Langs Film *Metropolis* (1927).

Freilich war die Beschäftigung mit Mächten, die der Mensch selbst geschaffen hat und die sich nach und nach seiner Kontrolle entziehen, lange Zeit ein rein theoretisches Unterfangen. Doch mit dem Aufblühen der Kybernetik nach dem Zweiten Weltkrieg begann sich das zu ändern. So führte beispielsweise Norbert Wiener im Mai 1949 in einem Seminarraum des Massachusetts Institute of Technology (MIT) einen dreirädrigen Roboter namens »Palomilla« vor, der sich selbsttätig auf eine Lichtquelle zubewegte. Damit gab es auch in der Praxis Beispiele, dass die Technik Anleihen bei kognitiven Leistungen von Lebewesen nehmen konnte, um Maschinen mit einer Art »Intelligenz« auszurüsten. Die alte Idee begann Realität zu werden – und damit wurden auch die Ängste davor immer realer. Eine Variante davon wird in Stanley Kubricks Film *2001: Odyssee im Weltraum* aus dem Jahr 1968 erzählt: An Bord des Raumschiffes »Discovery One« beginnt der Supercomputer HAL 9000, der mit einer künstlichen Intelligenz ausgestattet ist und das Raumschiff autonom steuern kann, ein Eigenleben zu entwickeln, die menschliche Besatzung als Gefahr zu betrachten und sie zu bekämpfen.

... in die Realität

Das Verhalten, das HAL 9000 im Film zeigt, entspricht dem, was von manchen Wissenschaftlern als »Superintelligenz« bezeichnet wird. Wikipedia definiert Superintelligenz als »Intellekt, der auch dem besten menschlichen Hirn in den meisten oder allen Bereichen überlegen ist, und zwar sowohl hinsichtlich kreativer und problemlöserischer Intelligenz wie auch bei sozialen Kompetenzen.«¹ Der Begriff tauchte erstmals im Jahr 1965 auf und wurde ab den späten 1990er-Jahren vor allem durch Nick Bostrom (University of Oxford) popularisiert.² Unterschieden wird dabei zwischen einer »schwachen« Superintelligenz – die qualitativ auf dem Niveau menschlicher Denkprozesse bleibt, aber quantitativ um ein Vielfaches schneller arbeitet – und einer »starken« Superintelligenz, die auf einem auch qualitativ übergeordneten Niveau operiert. Die Vertreter dieser Vorstellung lassen interessanterweise offen, ob der Superintelligenz auch Erinnerungsvermögen oder Selbstbewusstsein zukommt; ebenso nicht beantwortet wird, wie diese realisiert wird: biologisch, technisch oder als Hybrid dazwischen.

¹ <https://de.wikipedia.org/wiki/Superintelligenz> (18.04.2018).

² Nick Bostrom, »How Long Before Superintelligence«, in: *International Journal of Future Studies*, Bd. 2 (1998), S. 1–13.

From fantastic literature...

Figures escaping human control like the golem or the sorcerer's apprentice have always fired peoples' imagination and evoked dystopias—such as in Mary Shelley's *Frankenstein* from 1818 and countless other stories of nineteenth-century fantastic literature. With the rise of modern science, the technological element of these stories was given growing emphasis; more recently, they were also increasingly linked to political ideas. Just think of books like Aldous Huxley's *Brave New World* (1932), George Orwell's *1984* (1948), and Ray Bradbury's *Fahrenheit 451* (1953), or of the autocratic world of machines in Fritz Lang's film drama *Metropolis* (1927).

For a long time, the preoccupation with man-made forces gradually defying control had naturally remained a purely theoretical venture. This began to change, however, with the advent of cybernetics after World War II. In May 1949, for instance, in a seminar room at the Massachusetts Institute of Technology (MIT), Norbert Wiener presented a robot named "Palomilla" that was mounted on three wheels and automatically moved in the direction of a light source. Thus a precedent had been established that technology could indeed borrow cognitive achievements from humans in practice so as to furnish machines with a form of "intelligence." The old idea was becoming a reality—which implicated that the fears of what it would bring were likewise becoming more and more real. A variant of such a plot is told in Stanley Kubrick's film *2001: A Space Odyssey* from 1968: aboard the spaceship "Discovery One", the supercomputer HAL 9000, endowed with artificial intelligence and capable of autonomously steering the spaceship, starts developing a life of its own, identifying the human crew as a hazard and turning against it.

... to reality

The behavior HAL 9000 exhibits in the film resembles what several scientists refer to as "superintelligence." Wikipedia defines superintelligence as an "agent that possesses intelligence far surpassing that of the brightest and most gifted human minds [...] in most or virtually all domains of interest,"¹ both in terms of creative, problem-solving intelligence and social competence. The expression first appeared in 1965 and became popular above all through Nick Bostrom (University of Oxford) in the late 1990s.² One differentiates between "weak" superintelligence—which does not go beyond the quality level of human processes of thought, but as to quantity works many times faster—and "strong" superintelligence, also operating on a superior level in terms of quality. Interestingly enough, the proponents of this idea have left the question unanswered whether superintelligence also involves a capacity for remembering or a conscious mind; neither do they specify how it is brought about: biologically, technologically, or as a hybrid between those two.

¹ <https://en.wikipedia.org/wiki/Superintelligence> (April 18, 2018).

² Nick Bostrom, "How Long Before Superintelligence," *International Journal of Future Studies*, vol. 2 (1998), 1–13.

Eng mit Superintelligenz verwandt ist die »technologische Singularität«. Aus der Beobachtung heraus, dass die Leistungsfähigkeit von Computersystemen mit der Zeit exponentiell besser wird, haben sich Forscher in den späten 1950er-Jahren gefragt, wann denn der Zeitpunkt gekommen sein wird, in dem sich Maschinen mittels künstlicher Intelligenz selbst verbessern können. Dadurch, so die Idee, würde der technische Fortschritt dermaßen beschleunigt werden, dass die Zukunft der Menschheit nach diesem Ereignis nicht mehr vorhersehbar ist. Der Begriff Singularität bringt zum Ausdruck, dass der technische Wandel »so schnell und allumfassend ist, dass er einen Bruch in der Struktur der Geschichte der Menschheit darstellt«. Diese Formulierung stammt vom populärsten Vertreter dieser These, dem Futuristen (und Leiter der technischen Entwicklung bei Google) Raymond Kurzweil. 1999 hat er in seinem Buch *The Age of Spiritual Machines* (deutsch: *Homo S@piens*) prognostiziert, dass Computer um das Jahr 2030 die Menschen an Intelligenz übertreffen werden. Der prognostizierte Zeitpunkt der Singularität wurde allerdings schon mehrfach um Jahrzehnte in die Zukunft verschoben. Dennoch lebt die Idee weiter, dass dieser Sprung irgendwann eintritt.

Hinter Superintelligenz und Singularität versteckt sich ein weiterer zentraler Begriff: die »Intelligenzexplosion«, die 1965 von dem britischen Statistiker Irving John Good beschrieben wurde: »Eine ultraintelligente Maschine sei definiert als eine Maschine, die die intellektuellen Fähigkeiten jedes Menschen, und sei er noch so intelligent, bei Weitem übertreffen kann. Da der Bau ebensolcher Maschinen eine dieser intellektuellen Fähigkeiten ist, kann eine ultraintelligente Maschine noch bessere Maschinen bauen; zweifellos würde es dann zu einer explosionsartigen Entwicklung der Intelligenz kommen, und die menschliche Intelligenz würde weit dahinter zurückbleiben«, formulierte er. Nachsatz: »Die erste ultraintelligente Maschine ist also die letzte Erfindung, die der Mensch zu machen hat.«³ Ray Kurzweil stellte 2001 in dem Artikel »The Law of Accelerating Returns« die These auf, dass das Moore'sche Gesetz (laut dem sich die Rechenleistung von Computern ungefähr alle zwei Jahre verdoppelt) nur ein Spezialfall eines allgemeineren Gesetzes sei, nach dem die gesamte technologische Evolution verläuft: Das exponentielle Wachstum werde sich demnach auch in den Technologien fortsetzen, welche die heutigen Mikroprozessoren ablösen werden.⁴

Wie oben erwähnt, lassen die Vertreter der These der Superintelligenz offen, ob diese biologisch, technisch oder als Mischform realisiert werden kann. Letzteres ist die Basis des »Transhumanismus«: Durch den Einsatz technologischer Verfahren, also der Verschränkung von Biologie und Technologie, sollen die Grenzen menschlicher Möglichkeiten – intellektuell, physisch oder psychisch – erweitert werden. Der Begriff wurde in den späten 1960er-Jahren geprägt und u. a. von dem iranisch-amerikanischen Futuristen FM-2030 (geboren als F.M. Esfandiary) verbreitet: »Transhumane sind die erste Manifestation einer neuen Art von evolutionären Wesen. Sie ähneln darin den ersten Hominden, die vor vielen Millionen Jahren die Bäume verließen und begannen sich umzuschauen«, formulierte er 1989.⁵ Dieser Fortschrittsoptimismus stieß und stößt freilich auf harte Kritik. Bei genauer Betrachtung gibt es schon heute unzählige Cyborgs – etwa Träger von intelligenten Prothesen. Und manche

³ Irving John Good, »Speculations Concerning the First Ultraintelligent Machine« (1965), DOI: 10.1016/S0065-2458(08)60418-0, zitiert nach: https://de.wikipedia.org/wiki/Technologische_Singularit%C3%A4t (18.04.2018).

⁴ Ray Kurzweil, »The Law of Accelerating Returns«, <http://www.kurzweilai.net/the-law-of-accelerating-returns> (18.04.2018).

⁵ FM-2030, *Are You a Transhuman?*, New York 1989, S. 42.

“Technological singularity” is closely related to the idea of superintelligence. Due to their observation that the capacity of computer systems seemed to be increasing exponentially over time, scientists of the late 1950s wondered by when machines would be able to improve themselves through artificial intelligence. The idea was that this would accelerate technological progress to such an extent that the future of humanity would no longer be foreseeable after this event had occurred. The term singularity is meant to express the belief that technological change “is so rapid and so profound that it represents a rupture in the fabric of human history.” This came from the most popular proponent of the thesis, the futurist (and Google’s senior technology developer) Raymond Kurzweil. In 1999 he prophesied in his book *The Age of Spiritual Machines* that by around 2030 the intelligence of computers would be outperforming that of humans. In the meantime, though, the predicted date by when singularity will come about has been put off to the future by decades more than once. Nevertheless the idea prevails that one day this leap will take place.

Another central concept hiding behind superintelligence and singularity is that of “intelligence explosion,” which was described by the British statistician Irving John Good: “Let an ultra-intelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultra-intelligent machine could design even better machines; there would then unquestionably be an ‘intelligence explosion,’ and the intelligence of man would be left far behind,” he pointed out, adding: “Thus the first ultra-intelligent machine is the last invention that man need ever make.”³ In 2001, Ray Kurzweil, in his article “The Law of Accelerating Returns,” proposed the theory that Moore’s Law (according to which the processing power of computers doubles every year) was only a special case within a more universal law on which the entire technological evolution was based: exponential growth would therefore also continue in technologies about to replace today’s microprocessors.⁴

As mentioned above, the advocates of the theory of superintelligence fail to explain whether it can be realized biologically, technologically, or as a hybrid form. The idea of “transhumanism” is based on the latter: through the application of certain technological methods, i. e., the interlock of biology and technology, the limitations of human capacities are to be extended intellectually, physically, and psychologically. The term was coined in the late 1960s and proliferated thanks to such figures as the Iranian-American futurist FM-2030 (born as F.M. Esfandiary): “The transhuman is representative of the earliest manifestation of new evolutionary beings. They are like hominids who many millions of years ago came down from the trees and began to look around,” he verbalized in 1989.⁵ This optimistic belief in progress met and still meets with harsh criticism, of course. Taking a closer look, it turns out that countless cyborgs do already exist today—such as wearers of intelligent prosthetic limbs. Some scientists also regard smartphones and cloud-stored databases as extensions of human capacities through technology. Still, transhuman ideas continue to trigger a feeling of profound discomfort in many people, such

³ Irving John Good, “Speculations Concerning the First Ultraintelligent Machine” (1965), DOI: 10.1016/S0065-2458(08)60418-0, quoted from: <https://www.graphcore.ai/posts/ultraintelligence> (April 23, 2018).

⁴ Ray Kurzweil, “The Law of Accelerating Returns,” <http://www.kurzweilai.net/the-law-of-accelerating-returns> (April 18, 2018).

⁵ FM-2030, *Are You a Transhuman?*, New York 1989, 42.

Forscher sehen auch in Smartphones und Datenbanken in der Cloud eine Erweiterung der menschlichen Fähigkeiten durch Technologie. Dennoch lösen transhumane Ideen weiterhin großes Unbehagen bei vielen Menschen aus. So etwa zur Zeit bei den Plänen des us-Unternehmens Neuralink, das die Gehirne der Menschen über eine mit künstlicher Intelligenz angereicherte »Hirncoud« miteinander vernetzen will.⁶

Hintergrund der aktuellen Debatten

Diese vier Begriffe – die allesamt mit den alten Mythen von Golem, Zauberlehrling & Co verknüpft sind – bilden den Hintergrund für viele der aktuellen Debatten um künstliche Intelligenz. Sie spiegeln sich etwa in der Angst vor autonomen Waffensystemen wider, die ohne Zutun eines – verantwortlichen – Soldaten einen Menschen töten. Ein anderes Beispiel liefern Vorfälle im Frühling dieses Jahres: Im Zuge der Aufregung um Daten von zumindest 87 Millionen Nutzern des sozialen Netzwerkes Facebook, die offenbar im jüngsten Wahlkampf um die us-Präsidentschaft missbräuchlich Verwendung fanden, wurden wieder Befürchtungen wegen Datenspionage, Überwachung und Beeinflussung befeuert. Ein drittes aktuelles Beispiel ist das erste Todesopfer, das ein autonomes Automobil im März dieses Jahres in der Stadt Tempe in Arizona gefordert hat.

Diese Beispiele zeigen, dass die gesellschaftliche Diskussion, wie und unter welchen Umständen künstliche Intelligenz eingesetzt werden kann, darf und soll, noch lange nicht abgeschlossen ist – ja noch gar nicht richtig begonnen hat. Dennoch stellt sich schon jetzt die Frage, warum dieses Thema bei sehr vielen Menschen eher beängstigende Dystopien als positive Utopien hervorruft. Das gilt nicht für Laien, die die technischen Hintergründe und Beschränkungen nicht kennen, sondern für viele Insider wie Microsoft-Gründer Bill Gates, Tesla-Chef Elon Musk oder den heuer verstorbenen Astrophysiker Stephen Hawking, denen die schnelle Entwicklung der künstlichen Intelligenz nicht geheuer ist.⁷

Existenzielle Kränkungen des Menschen

Als mögliche Erklärung wird immer wieder angeführt, dass künstliche Intelligenz eine weitere »existenzielle Kränkung« des Menschen darstelle. Angelehnt ist dieser Begriff an die drei Kränkungen der Menschheit, die Sigmund Freud 1917 postuliert hat. Diese waren jeweils eine Konsequenz moderner Wissenschaft und haben das Selbstverständnis der Menschen in Frage gestellt: erstens die »kosmologische Kränkung« nach der Erkenntnis von Nikolaus Kopernikus, dass die Erde nicht der Mittelpunkt des Weltalls ist; zweitens die »biologische Kränkung« nach der Entdeckung von Charles Darwin, dass der Mensch aus einer Tierreihe hervorgegangen ist; und drittens die »psychologische Kränkung«, weil sich ein Teil des Seelenlebens dem bewussten Willen entzieht.⁸ Nun wird auch der Geist als bisher eifersüchtig bewachtes Alleinstellungsmerkmal des Menschen in Frage gestellt. ✕

⁶ Miriam Meckel, »Brainhacking. Der Spion in meinem Kopf«, in: ZEIT ONLINE, 11.04.2018, <http://www.zeit.de/2018/16/brainhacking-gehirn-kopf-konzerne-miriam-meckel> (18.04.2018).

⁷ »Hawking: AI ist das Beste für die Menschen – oder das Schlechteste« (26.10.2016), <https://www.computerwoche.de/a/hawking-ai-ist-das-beste-fuer-die-menschen-oder-das-schlechteste,3325701> (18.04.2018).

⁸ [https://de.wikipedia.org/wiki/Kränkungen_der_Menschheit](https://de.wikipedia.org/wiki/Kr%C3%A4nkungen_der_Menschheit) (18.04.2018).

as when currently faced with the plans of the us company Neuralink, which seeks to link human brains via a “brain cloud” enhanced with artificial intelligence.⁶

⁶ Miriam Meckel, “Brainhacking. Der Spion in meinem Kopf,” ZEIT ONLINE, April 11, 2018, <http://www.zeit.de/2018/16/brainhacking-gehirn-kopf-konzerne-miriam-meckel> (April 18, 2018).

In the background of current debates

These four concepts—all of which are related to the old myths of the golem, the sorcerer’s apprentice, and the like—linger in the background of many a current debate about artificial intelligence. They are, for instance, reflected in the dread of autonomous weapon systems capable of killing people without a—responsible—soldier involved. Incidents arising in the spring of this year provide another example: the turmoil caused by the apparent abuse of the data of at least 87 million users of the social network Facebook for the recent us presidential election campaign once again fueled fears of data espionage, surveillance, and manipulation. A third topical example is the first casualty claimed by an autonomous automobile in the city of Tempe in Arizona in March this year.

These examples illustrate that the discussion in society about how and under what circumstances artificial intelligence can, may, or should be applied will continue for a long time to come—and has actually only just begun. Nevertheless, the question arises why the subject evokes horrifying dystopias rather than optimistic utopias in so many people. We are not talking about non-specialists unfamiliar with the technological backgrounds and limits, but about numerous insiders, including the founder of Microsoft Bill Gates, Tesla boss Elon Musk, and the recently deceased astrophysicist Stephen Hawking, all of whom feel uncomfortable with the rapid development of artificial intelligence.⁷

⁷ “Hawking: AI ist das Beste für die Menschen—oder das Schlechteste” (October 26, 2016), <https://www.computerwoche.de/a/hawking-ai-ist-das-beste-fuer-die-menschen-oder-das-schlechteste>, 3325701 (April 18, 2018).

Humanity’s existential injuries

As a possible explanation for this it has repeatedly been stated that artificial intelligence might represent a further “existential injury” inflicted upon humankind. The term alludes to the three injuries to man postulated by Sigmund Freud in 1917. Each of them was the result of modern science and questioned man’s self-conception: first, the “cosmological injury” resulting from Nicolaus Copernicus’s discovery that the Earth was not at the center of the universe; second, the “biological injury” caused by Charles Darwin’s discovery that man descended from animal ancestors; and third, the “psychological injury,” as man’s inner life in part refuses to be controlled by conscious will.⁸ Now the intellect, too, has been cast into doubt as man’s jealously guarded unique selling proposition. ✕

⁸ https://de.wikipedia.org/wiki/Kränkungen_der_Menschheit (April 18, 2018).

Helmut Leopold

Ist KI eine sichere Technologie, und können wir ihr vertrauen?



Der zunehmende Einsatz künstlicher Intelligenz wirft zahlreiche Sicherheitsfragen auf. Zum einen ermöglicht die Technologie eine Ausweitung herkömmlicher Formen der Kriminalität, zum anderen treten aber auch völlig neue Formen der Cyberbedrohung auf – zu deren Abwehr es erst wenige Methoden gibt. Manche Befürchtungen erscheinen allerdings angesichts des derzeitigen Entwicklungsstandes von künstlicher Intelligenz deutlich übertrieben.

Eines der grundlegenden Probleme von Systemen künstlicher Intelligenz (KI) ist die Frage des Vertrauens in deren korrekte Funktion. Für herkömmliche Systementwicklungen gibt es seit Jahrzehnten bewährte Testmethoden und Zertifizierungsprozesse. Im Wesentlichen beruhen alle diese Methoden darauf, dass Ingenieure sowohl die Systemdesigns als auch die Entwicklungsprozesse genau kennen. Auf dieser Basis können Prüfvorschriften und Testprozesse festgelegt werden, um die Korrektheit von Systemfunktionen zu verifizieren, und Systemverhalten und Entscheidungsfindungen können von Experten nachvollzogen und erklärt werden.

Bei KI-Systemen wird es für Experten zunehmend schwieriger, die Entscheidungsfindung nachzuvollziehen. Bei neuronalen Netzen lässt sich nicht mehr feststellen, aufgrund welcher Parameter von der Maschine eine bestimmte Entscheidung getroffen wurde. Wir müssen annehmen und hoffen, dass richtig gelernt wurde, und können nur darauf vertrauen, dass das Ergebnis richtig ist. Es gibt keine Möglichkeit, Algorithmen zu kontrollieren, wie es bisher bei deterministisch gebauten technischen Systemen möglich war.

Täuschung der KI-Algorithmen

Dieser Umstand kann durchaus zu grundlegenden Problemen führen, wie bereits diverse Beispiele zeigen. Ein bekanntes Beispiel ist der jüngste Unfall von Tesla.¹ Erste Forschungsarbeiten belegen, wie KI-Algorithmen durch einfache physische Veränderungen der zu identifizierenden Objekte, wie etwa durch Graffiti auf Verkehrsschildern, leicht durcheinanderkommen und schließlich in ihrem Urteilsvermögen getäuscht werden. Obwohl solche Veränderungen für den Menschen leicht erkennbar sind, haben künstliche Maschinen damit sehr zu kämpfen. Durch den gezielten Einsatz solcher »Robust Physical Perturbations«, etwa in Form einfacher Papierausdrucke, können KI-Kamerasysteme von Autos völlig getäuscht werden.²

Dies kann bei digitalen Bildern noch einen Schritt weitergehen. So kann etwa durch die für Menschen unsichtbaren Veränderungen nur einiger Pixel ein KI-Algorithmus so getäuscht werden, dass statt eines Schulbusses ein

Seit Jänner 2009 leitet Helmut Leopold am Austrian Institute of Technology (AIT) das Center Digital Safety & Security. In dieser Rolle ist er für das Sicherheitsforschungsprogramm des AIT und damit für Cybersecurity und -schutz kritischer IT-Infrastrukturen zuständig. Zuvor war er bei Telekom Austria in verschiedenen Managementfunktionen tätig und als Technologiechef der Telekom Austria maßgeblich für die digitale Transformation des Unternehmens zu einem modernen Multimedia- und Breitbandunternehmen verantwortlich. Helmut Leopold ist derzeit auch Präsident der Gesellschaft für Informations- und Kommunikationstechnologien (GIT) im Österreichischen Verband für Elektrotechnik (ÖVE) sowie Mitglied des ÖVE-Vorstandes.

¹ Siehe »Die können was! Aber können Roboter auch fühlen? Raúl Rojas und Thea Dorn über künstliche Intelligenz – Doppelinterview«, in: *Aus Politik und Zeitgeschichte. Künstliche Intelligenz*, 68. Jahrgang (2018), Nr. 6–8; <http://www.bpb.de/apuz/263675/die-koennen-was-aber-koennen-roboter-auch-fuehlen> (14.05.2018).

² Siehe Kevin Eykholt u. a., »Robust Physical-World Attacks on Deep Learning Visual Classi-



Since January 2009, Helmut Leopold has been with the Austrian Institute of Technology (AIT) where he is Head of the Digital Safety & Security Center. In this role he is responsible for the ICT-based research program focusing on cybersecurity, critical infrastructure protection, and public safety. Before AIT, he was with Telekom Austria, the largest network operator in Austria, in various management positions. He played a leading role in the innovation process of the company and was responsible for the digitization and next generation network transformation program of the IT and network infrastructure. Helmut Leopold is President of the Austrian Organization for Information and Communication Technology (GIT) and Board Member of the Austrian Electro-technical Association (ÖVE).

Helmut Leopold

Is AI a Safe Technology, and Can We Trust It?

Increasing deployment of artificial intelligence raises a number of safety and security issues. For one thing, the technology facilitates the expansion of traditional types of criminal activity. For another, there also are entirely new types of cyberthreats emerging—and only few methods available to defend against them. Some fears, however, seem clearly overstated given the current state of development in the field of artificial intelligence.

One fundamental problem of artificial intelligence (AI) systems is the question of how much trust to place in their correct functioning. For conventional system developments, proven testing methods and certification processes have been in place for decades. In essence, all such methods are based on engineers' in-depth knowledge of both system designs and development processes. On this basis, test instructions and processes can be defined to verify the functional correctness of the system, and system behaviors and decision-making processes can be traced and analyzed by experts.

In AI systems, experts are having increasing difficulties to trace decision-making processes. In neural networks, it can no longer be determined based on what parameters the machine has come to make a certain decision. We must assume and hope that things were learnt correctly and can only trust that the results are correct. There is no possibility of controlling algorithms the way it was previously possible with deterministically constructed technological systems.

Deceiving AI algorithms

This is a fact that may well lead to elemental problems, as several examples have already shown. One well-known case is the most recent Tesla accident.¹ First research projects substantiate how easily AI algorithms may be perturbed by simple physical modifications, such as graffiti on traffic signs, and eventually deceived in their judgement. While such modifications are easily identifiable for humans, artificial machines are struggling hard with them. Through purposive use of such "Robust Physical Perturbations," e. g. in the form of simple printouts on paper, AI camera systems of cars can be totally deceived.²

In digital images, this may even go one step further. So, for example, an AI algorithm can, through modification of just a few pixels, invisible for humans, be so deceived that what is a school bus is instead identified

¹ See "Die können was! Aber können Roboter auch fühlen? Raúl Rojas und Thea Dorn über künstliche Intelligenz – Doppelinterview," *Aus Politik und Zeitgeschichte. Künstliche Intelligenz*, vol. 68 (2018), nos. 6–8; <http://www.bpb.de/apuz/263675/die-koennen-was-aber-koennen-roboter-auch-fuehlen> (May 14, 2018).

Strauß erkannt wird.³ Das Bedrohliche ist, dass sich nicht genau feststellen lässt, aufgrund welcher Parameter solche Fehlinterpretationen erfolgen. Daher liegt ein Schlüssel für den erfolgreichen Einsatz der KI-Technologie in der Verfügbarkeit von ausreichend großen Mengen an Lerndaten, die möglichst alle Kombinationen von Lernzuständen abbilden und von Experten bewertet und beurteilt worden sind.

Auch KI ist abergläubisch

Burrhus F. Skinner, der Pionier der Verhaltensforschung und des programmierten Lernens, fand bereits 1948 heraus, dass Tauben abergläubisch sind.⁴ In einem Versuch stellte er fest, dass zufällige Zusammenhänge zwischen Ereignissen, zum Beispiel zwischen zur Verfügung gestelltem Futter und zufälliger Körperhaltung, als Sollmuster erlernt werden. Solche Phänomene sind auch bei KI-Algorithmen zu beachten.

Zwei grundlegende Probleme von KI-Systemen bestehen in der Schwierigkeit der Modellbildung und in der Anzahl der zur Beurteilung betrachteten Parameter. Wird das Modell zu kompliziert gewählt und sind zu viele Parameter zu beachten, so besteht die Gefahr von »Overfitting«. Die Maschine ist dann nicht mehr in der Lage, unnötige Information (»noise«) von echter Information (»signal«) zu unterscheiden. Der KI-Algorithmus lernt, salopp ausgedrückt, die Problemstellung aus den nichtrelevanten Teilen der Lerndaten auswendig, ohne die eigentliche Aufgabe zu lösen.

Diese Problematik kennen wir auch aus herkömmlichen Betrachtungen und Versuchen, korrekte Vorhersagen zu treffen – etwa in Wirtschaftssystemen, bei Wetter und Klima, Sport, Spielen etc.⁵

Unerwünschte Backdoors

Da KI-Systeme auf der Verfügbarkeit geeigneter Lerndaten beruhen, um von diesen bestimmte Muster einzutrainieren, die dann in der echten Welt angewandt werden, können dadurch natürlich auch bestimmte Mechanismen, sogenannte Backdoor-Funktionen, unerkannt mittrainiert werden. So werden zum Beispiel unter Umständen bestimmte Gesichtsbilder mittrainiert, sodass es bei einer Passkontrolle bestimmter Personen in der Realität zu einem verfälschten Ergebnis kommt. Es gibt keine Möglichkeit, im Nachhinein ein solches verstecktes Verhalten in einer KI-Maschine zu identifizieren, weil das Geheimnis nur in den verwendeten Lerndaten vorhanden und im KI-System nicht mehr ersichtlich und nachvollziehbar ist.

Somit kommen wir zur grundlegenden Frage, woher die Lerndaten kommen, wer diese verifiziert hat und wie Lernprozesse von KI-Algorithmen vor sich gehen. Open-Source-Ansätze sind sicher eine vielversprechende Möglichkeit, Vertrauen in KI-Systeme zu fördern. Alternative Ansätze sind Zertifizierungen von Lerndaten, aber dafür müssten Mechanismen definiert werden, wie Zertifizierungen erfolgen sollen.

KI wird Cyberkriminalität pushen

Wie bei jeder Einführung einer neuen Technologie in eine Gesellschaft stellen sich nicht nur die eigentlich beabsichtigten positiven Vorteile der neuen Funktionen ein, sondern sind diese auch Missbrauch ausgesetzt. Missbrauch

fiction«, CVPR 2018, 10.04.2018; <https://arxiv.org/pdf/1707.08945.pdf> (14.05.2018). Siehe auch Mark Samuels, »Hacking Risk for Computer Vision Systems in Autonomous Cars«, in: *SecurityIntelligence*, 10.08.2017; <https://securityintelligence.com/news/hacking-risk-for-computer-vision-systems-in-autonomous-cars/> (14.05.2018).

³ Siehe Dave Gershgorin, »Fooling The Machine. The Byzantine science of deceiving artificial intelligence«, in: *Popular Science*, 30.03.2016; <https://www.popsci.com/byzantine-science-deceiving-artificial-intelligence/> (14.05.2018).

⁴ Siehe Burrhus F. Skinner, »Superstition in the Pigeon«, Indiana University, 05.06.1947; vgl. Michael Niedeggen, »Skinner und seine abergläubischen Tauben«, Seminar »Klassische Experimente der Psychologie«, http://www.ewi-psy.fu-berlin.de/einrichtungen/arbeitsbereiche/allgpsy/media/media_lehre/Lernen_und_Ged__cht__s/seminar_6.pdf (14.05.2018).

⁵ Siehe Nate Silver, *the signal and the noise: why so many predictions fail—but some don't*, New York u. a. 2012.

² See Kevin Eykholt a.o., “Robust Physical-World Attacks on Deep Learning Visual Classification,” CVPR 2018, April 10, 2018; <https://arxiv.org/pdf/1707.08945.pdf> (May 14, 2018). See also Mark Samuels, “Hacking Risk for Computer Vision Systems in Autonomous Cars,” *SecurityIntelligence*, August 10, 2017; <https://securityintelligence.com/news/hacking-risk-for-computer-vision-systems-in-autonomous-cars/> (May 14, 2018).

³ See Dave Gershgorn, “Fooing The Machine. The Byzantine science of deceiving artificial intelligence,” *Popular Science*, March 30, 2016; <https://www.popsci.com/byzantine-science-deceiving-artificial-intelligence> (May 14, 2018).

⁴ See Burrhus F. Skinner, “Superstition in the Pigeon,” *Indiana University*, June 5, 1947; cf. Michael Niedeggen, “Skinner und seine abergläubischen Tauben,” seminar “Klassische Experimente der Psychologie”; http://www.ewi-psy.fu-berlin.de/einrichtungen/arbeitsbereiche/allgpsy/media/media_lehre/Lernen_und_Ged__chtn__s/seminar_6.pdf (May 14, 2018).

⁵ See Nate Silver, *the signal and the noise: why so many predictions fail—but some don't* (New York, a.o., 2012).

as an ostrich.³ The ominous thing is that it cannot be precisely determined what parameters lead to such misinterpretations. Hence one key to successful implementation of AI technology lies in the availability of sufficient amounts of learning data, which represent, if possible, all potential combinations of learning states and have been evaluated and assessed by experts.

AI is superstitious, too

Burrhus F. Skinner, the pioneer of behavioral science and programmed learning, found out already in 1948 that pigeons are superstitious.⁴ In an experiment he observed that coincidental connections between events, like, for example, the provision of birdseed and body posture, are learnt as desired action patterns. Such phenomena are also observable in AI algorithms.

Two fundamental problems of AI systems are the difficulty of model formation and the number of parameters to be considered. If the model is made too complicated and too many parameters are to be taken into account, there is a danger of “overfitting.” The machine will then be no longer able to distinguish between unnecessary (“noise”) and real information (“signal”). In non-technical terms, the AI algorithm learns the problem by rote from non-relevant segments of the learning data without solving the actual task.

This is a problem that we are also familiar with from conventional analyses and attempts of making correct predictions—for example, in economic systems, weather and climate forecasts, sports, games, etc.⁵

Unwanted backdoors

As AI systems rely on the availability of suitable learning data to train specific patterns from them, which are then applied in the real world, there is of course a possibility that certain mechanisms, so-called backdoor functions, are trained as well without being identified. So, for example, certain facial patterns may be trained, which, in passport controls of certain people, may lead to falsified results. There is no possibility to identify such hidden behavior in an AI machine after the event, because the secret is only included in the learning data used and is no longer identifiable, nor reproducible, in the AI system itself.

This brings us to the fundamental question of where the learning data come from, who has verified them, and how learning processes of AI algorithms occur. Open-source approaches certainly are a promising possibility of furthering trust in AI systems. An alternative approach is the certification of learning data, but this would require defining mechanisms of how to conduct certification processes.

AI will boost cybercrime

Like with any new technology introduced in a society, it is not only the intended positive advantages of the new functions that are making themselves felt, but these are also subject to abuse. Abuse has all along been a fixed element of innovative implementations that must always be

ist seit jeher ein fester Bestandteil innovativer Implementierungen, den es immer mit zu betrachten gilt. Die starke Ausbreitung von KI-Technologien wird auch zu neuen Formen der Cyberkriminalität, neuen Möglichkeiten für Terrorismus und zu politischen Störungen führen. Es werden neuartige Bedrohungen entstehen, und der bislang typische Charakter existierender Bedrohungen für kritische Infrastrukturen wird sich ändern. Deshalb raten Experten zur schnellen Implementierung eines regulativen Rahmens, um die missbräuchliche Verwendung von KI zu verhindern.⁶

Bisher konzentriert man sich nur auf die möglichen positiven Aspekte von KI in den verschiedenen Anwendungsdomänen. Es gibt kaum Initiativen, die sich mit der Vermeidung von Missbrauch der neuen Technologie beschäftigen. Wenn die Gesellschaft nicht schnell genug etwas gegen diese heraufdämmern- den Gefahren unternimmt, bekommen kriminelle Akteure mit KI neue effektive Werkzeuge für ihre Ziele in die Hand. Der Bogen potenziellen Missbrauchs spannt sich von herkömmlicher Kriminalität über den gesellschaftlichen bis zum staatlichen Bereich. Wenn Lerndaten wie oben beschrieben schlecht aus- gesucht oder manipuliert werden, können Kriminelle diverse Anwendungen für ihre Zwecke missbrauchen.

KI kann das grundsätzliche Problem der Manipulation öffentlicher Mei- nungen in der neuen Medienwelt wesentlich verschärfen. Durch personali- sierte Desinformationskampagnen lässt sich jedes Individuum mit einem aus- gewählten Set an Informationen in seinem (Wahl-)Verhalten beeinflussen. KI- Systeme bieten effektive Möglichkeiten, Gesinnungen und politische Einstellun- gen aus verfügbaren Daten der sozialen Onlinemedien oder anderen digitalen Spuren zu ermitteln. KI ist ein Werkzeug, das für autokratische und diktator- ische Systeme geradezu zu einem Wundermittel werden kann.

Noch besteht jedoch die Möglichkeit, eine Überlegenheit für die »gute« Nutzung von KI herzustellen. Es müssen unbedingt Forschungsinitiativen ge- setzt werden, um die durch KI-Technologien neuen Bedrohungslagen zu erken- nen und zu verstehen und geeignete Verteidigungsmaßnahmen zu entwickeln.

Es gilt sich die Frage zu stellen, wie Software und Hardware in Zukunft designt werden sollen und welche Gesetze und Regulative vor dem Hintergrund digitaler Hochrüstung unbedingt eingeführt werden müssen. Es ist notwendig, eine Balance zwischen der Offenheit von KI und proprietären Systemen zu erreichen und Mechanismen zu finden, welche die Kontrolle von KI-Systemen ermöglichen. Die politischen Rahmenbedingungen müssen an die neuen KI-Herausforderungen angepasst werden.

Wird KI die Menschheit übertreffen?

Der amerikanische Tech-Guru und Zukunftsseher Ray Kurzweil, der mit Unterstützung von Google und der NASA den Silicon-Valley-Thinktank »Singularity University« und damit eine weltweite Bewegung für die Entstehung der posthumanen Gesellschaft ins Leben gerufen hat, in der sich mit Cyborgs eine Art Zwitterwesen aus Menschen und Maschinen tummeln, sieht die Ablöse menschlicher Intelligenz durch KI im Jahr 2047 gekommen. Bei seiner expo- nentiellen Hochrechnung stützt er sich auf das Mooresche Gesetz von 1965, demzufolge sich die Leistung von Computerchips alle zwei Jahre verdop- peln wird – eine Vorhersage, die sich als zutreffend erwiesen hat.

⁶ Siehe Miles Brundage u. a., *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*, Future of Humanity Institute u. a. 2018; <https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf> (14.05.2018).

considered. The wide spread of AI technologies will also lead to new forms of cybercrime, new possibilities for terrorism, and political interferences. New threats will emerge, and the hitherto typical character of existing threats to critical infrastructures will change. Experts therefore recommend the quick putting in place of a regulatory framework in order to prevent AI abuse.⁶

⁶ See Miles Brundage a.o., *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation* (Future of Humanity Institute, a.o., 2018); <https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf> (May 14, 2018).

Until now, the focus has only been on the potential positive aspects of AI in different areas of application, and there are hardly any initiatives that address the prevention of abuse of the new technology. Unless society takes swift action against these emergent dangers, criminal actors using AI will be able to get their hands on new and highly effective tools to pursue their goals. The range of potential abuse extends from traditional crime to the areas of society and government at large. If learning data are wrongly selected or manipulated as described above, criminals may be able to abuse various different applications for their purposes.

AI may also dramatically worsen the fundamental problem of the manipulation of public opinion in the new media world. Targeted personalized disinformation campaigns make it possible to influence any individual's (voting) behavior with sets of selective information. AI systems of course offer effective possibilities of filtering out attitudes and political views from available online social media data and other digital traces. AI thus is a tool that may turn into a kind of magic wand for autocratic and dictatorial regimes.

There still is a window of opportunity to make sure that the “good” use of AI prevails. Unquestionably, research initiatives must be launched to identify and understand the new threats ushered in by AI technologies and to develop adequate defense measures.

The question must be raised of how software and hardware will have to be designed in the future and what laws and regulations must indubitably be introduced against the backdrop of the digital buildup. It will be necessary to strike a balance between the openness of AI and proprietary systems and to find mechanisms that facilitate control of AI systems. The framework of political conditions must be adapted so as to be up to the new challenges posed by AI.

Will AI leave humankind behind?

American tech guru and futurist Ray Kurzweil, who, with the help of Google and NASA, has started the Silicon Valley think tank “Singularity University” and, with it, a global movement for the establishment of a post-human society populated by cyborgs, a kind of hybrid between human and machine, sees the supersession of human intelligence by AI coming in in the year 2047. For his exponential projection, he relies on Moore’s law of 1965, according to which the performance of computer chips will double every two years—a prediction that has proven correct.

It should, however, be noted that in all forecasts made about AI so far eventual developments did not match the time horizon. In this context, and in view of the current AI euphoria, the Moravecian paradox⁷ should

⁷ See https://en.wikipedia.org/wiki/Moravec%27s_paradox (May 14, 2018).

Allerdings gilt es zu beachten, dass bei allen bisherigen Prognosen über KI der Zeithorizont nicht der tatsächlichen Entwicklung entsprach. In diesem Zusammenhang ist angesichts der Euphorie für KI-Leistungen an das Moravec'sche Paradox⁷ zu erinnern: Scheinbar komplizierte Abläufe (z. B. Schachzüge) sind leichter durch KI berechenbar als einfache Abläufe wie die unseres Alltags. Wenn es um Wahrnehmung und Bewegung geht, legt ein Kleinkind eine höhere Komplexität an den Tag als jedes erdenkliche KI-System.

Wichtige menschliche Eigenschaften, die erforderlich sind, um komplexe Herausforderungen des Lebens zu meistern, sind Flexibilität, Fantasie, Kreativität etc. – was üblicherweise als Eingebung, Erfindung und Lösungsfähigkeit beschrieben wird. Das menschliche Leben ist keine lineare Entwicklung, sondern stets durch disruptive Effekte bestimmt, die durch Nichtlinearitäten, Fehler und Zufälle beeinflusst werden. Diese Mechanismen werden durch Maschinen noch länger nicht nachahmbar sein.

Dieses lineare Verständnis ist einer der größten Irrtümer der Big-Data- und KI-Disziplin, dem man heute schon im Internet begegnet. Nur Angebote zu bekommen, die unserem Kaufverhalten oder unseren Vorlieben entsprechen, scheint zwar auf den ersten Blick »intelligent« und »hilfreich« zu sein. Aber diese sind nur die eine Hälfte des wahren Lebens. Das Leben ist kein steriler Platz der Abarbeitung von Fakten – es macht Spaß, ist spannend, herausfordernd und manchmal auch scheinbar »sinnlos«. Und am allerwichtigsten: Der Zufall spielt eine sehr große und oft bestimmende Rolle. Wahre Innovation entsteht durch Kreativität, Zufall und Kombination von scheinbar nicht zusammenpassenden Dingen. Im Prinzip kann man sogar sagen, dass KI Kreativität vernichten kann.

Wir wissen heute noch viel zu wenig über den Bauplan des menschlichen Gehirns, als dass wir in der Lage wären, dieses bioinformatische Wunderwerk zu entschlüsseln und Maschinen nach seinem Ebenbild zu formen. Doch wir müssen uns schon heute eingehend damit beschäftigen, wie wir in Zukunft mit der massenhaften Delegation gesellschaftlicher Aufgaben an maschinelle Intelligenz umgehen wollen.

Die ethische Dimension von KI

Die rasanten technologischen Fortschritte der KI werfen auch grundsätzliche gesellschaftliche Fragen auf. KI hat einerseits enormes Potenzial zur Verbesserung der Welt, andererseits bleibt jedoch das Faktum bestehen, dass auch die kühnste Maschine den Menschen in seiner Mehrdimensionalität aus Kognition, Emotion und Körperlichkeit und in seiner sinnlichen Wahrnehmung der Welt niemals wird erreichen können.

Die Maschine ist in vielem besser und schneller als der Mensch und muss dort für die Bewältigung der großen anstehenden ökonomischen, ökologischen und sozialen Fragen eingesetzt werden, wo sie gebraucht wird. Für menschlichen Forscherdrang wird es niemals eine Grenze geben – das ist Grund zur Hoffnung, besonders im Hinblick auf die immer größer werdenden Probleme der globalen, digital vernetzten Gesellschaft von heute.

Die hier noch fehlenden Einsichten und das Nichtverstehen machen für Größenwahn anfällig. Wie aus einer kindlichen Trotzreaktion heraus soll der künstliche Mensch erschaffen werden. Es bleibt aber ein großer Unterschied, ob eine Maschine auf die Wahrnehmung von emotionalen Stimmungen trainiert

⁷ Siehe https://en.wikipedia.org/wiki/Moravec%27s_paradox (14.05.2018).

be brought to mind: apparently complex processes (such as chess moves) are more easily computable by AI than simple human doings like those of our everyday life. When it comes to perception and mobility, a baby shows greater complexity than any conceivable AI system.

Important human qualities that are needed to cope with the challenges and complexities of life are flexibility, imagination, creativity, etc.—which are usually described as inspiration, invention, and problem-solving skill. Human life is not a linear development but defined by disruptive effects, which are influenced by nonlinearities, mistakes, and accidents. Such mechanisms will still not be imitable by machines any time soon.

Such linear understanding is one of the biggest misconceptions in Big Data and AI that one already comes across on the Internet today. Being only shown offers that are in line with one's purchasing history or preferences may indeed seem "intelligent" or "helpful" at first sight. But this is just one half of what life is all about. Life is not a sterile site of processing facts—it's fun, exciting, challenging, and sometimes also apparently "senseless." And most importantly, accident plays a very large and often defining role. True innovation is the result of creativity, accident, and the combination of seemingly incongruous things. In general, one might even say that AI can destroy creativity.

We still know far too little about the structural blueprint of the human brain to be able to decode the workings of this miracle of computational biology and to build machines in its image. However, already today we have to intensively concern ourselves with how to deal with the mass delegation of social tasks to machine intelligence in the future.

The ethical dimension of AI

The rapid progress of AI technology also raises fundamental social questions. On the one hand, AI has enormous potential for improvement of the world; on the other, the fact remains that even the boldest machine will never come equal to the human in its multidimensionality of cognition, emotion, and corporeality, and its sensory perception of the world.

A machine can do many things better and faster than a human and must be put to good use wherever it is needed in coping with the big imminent economic, ecological and social issues. There will never be a limit to the inquiring human mind—this is reason for hope, particularly so with regard to the growing problems of today's global, digitally networked society.

Lack of insight and full understanding of the matter may cause megalomania. As if out of infantile defiance, man wants to create the artificial human. Yet it still does make a big difference whether a machine can be trained to perceive and respond to emotional moods or starts having its own. Only the human has a body-and-soul configuration, and also has, as aside from cognitive capacities to transform the world, an experiential level in the form of emotions. Only grief, joy, love, empathy, sympathy, compassion, or readiness to help provide humans with the unique range that distinguishes them from everything else.

niert werden kann oder ob sie selbst welche hat. Nur der Mensch hat eine Körper-Seele-Konstellation, hat neben seinen kognitiven Fähigkeiten zur Gestaltung der Welt auch eine Erlebnisebene in Form von Gefühlen. Erst Trauer, Freude, Liebe, Empathie, Mitleid oder Hilfsbereitschaft geben dem Humanen jene einzigartige Spannung, die uns von allem unterscheidet.

Daher muss gerade im Angesicht von KI die ethische Richtung diskutiert werden, die in Zukunft eingeschlagen werden soll. Will man nicht, dass eines Tages Algorithmen regieren, müssen Menschen die Endverantwortung für das Tun der Maschinen behalten. Der Wert des Lebens bestimmt sich nicht nach rein wirtschaftlichen Gesichtspunkten, lässt sich nicht nach mathematischen Formeln festlegen. Daher darf KI keine neue Religion werden.

Die Europäische Kommission hat das erkannt und wird auf Grundlage der »Erklärung der Europäischen Gruppe für Ethik in der Wissenschaft und Neuen Technologien« vom 9. März 2018⁸ eine Expertengruppe einsetzen, die in einer breiten Allianz unterschiedlicher Interessenvertreter aus Politik, Wirtschaft und Wissenschaft Ethikleitlinien für KI entwickeln soll.

Europäische Grundwerte sollen in diesem Entwicklungsprozess berücksichtigt werden. Entscheidend ist das nicht zuletzt für die aktuelle europäische Initiative zur Förderung der Entwicklung von Technologien künstlicher Intelligenz vom 8. April 2018, mit der die F&E-Investitionen im KI-Bereich bis 2020 um mindestens 20 Milliarden Euro erhöht werden sollen.⁹ Die Kernfragen, die dabei zu lösen sein werden, betreffen die Einhaltung von EU-Grundrechten wie Fairness, Sicherheit, Transparenz und Solidarität, aber auch moralische Regelungen für die Zukunft der Arbeit und für die Sicherung des Fortbestandes der Demokratie.

Emotionale Intelligenz zur Beherrschung der Technik

Friedrich Dürrenmatts Theaterstück *Die Physiker*, das die grundlegende Frage der Verantwortung für technologische Entwicklungen aufwirft, welche die Gefahr einer Bedrohung der Welt mit sich bringen, gewinnt so neue Aktualität. Technikbeherrschung ist ein unbedingtes Muss aller menschlicher Bestrebungen – die Essenz des Menschseins kreist doch um Freiheit, Willen und Kreativität.

Es stellt sich die Frage, ob die derzeitigen Ausbildungssysteme diesen neuen Herausforderungen Rechnung tragen können. Die heutigen Aufnahmekriterien für Universitäten und Fachhochschulen setzen zurzeit falsche Prioritäten. In vielen Disziplinen wird nach wie vor umfangreichem Faktenwissen der Vorzug gegeben; Empathie, Kreativität und Lösungskompetenz bei unvollständigen und widersprüchlichen Faktenlagen sowie gesamtheitlichen Betrachtungen von komplexen Systemen wird zu wenig Augenmerk geschenkt.

Es gilt das Ausbildungssystem grundlegend zu überdenken. Um mit Robotern und KI in Zukunft leben und die Technik beherrschen zu können, wird es notwendig sein, sich auf die menschlichen Kernkompetenzen der emotionalen Intelligenz und wahren Bildung zu besinnen und triviales Faktenlernen durch Soft Skills zu ersetzen. Man kann nicht mit KI und Robotern in Wettbewerb treten, sondern muss sie beherrschen. ✕

⁸ Siehe »Künstliche Intelligenz: Die Europäische Kommission beginnt um Ethikstandards und modernste Technik zusammen zu bringen«, Pressemitteilung der Europäischen Kommission, 09.03.2018; http://europa.eu/rapid/press-release_IP-18-1381_de.htm (14.05.2018).

⁹ Siehe »Künstliche Intelligenz: Kommission beschreibt europäisches Konzept zur Förderung von Investitionen und Entwicklung ethischer Leitlinien«, Pressemitteilung der Europäischen Kommission, 25.04.2018; www.europa.eu/rapid/press-release_IP-18-3362_de.pdf (14.05.2018).

We must therefore, especially in the face of AI, discuss the path we want to take in the future. If we do not want to be governed by algorithms someday, we must stay in ultimate control of the workings of machines. The value of human life is not determined by purely economic aspects and cannot be defined by mathematical formulae. AI must therefore not become a new religion.

The European Commission is aware of this and will, on the basis of the “Declaration of the European Group on Ethics in Science and New Technologies” of March 9, 2018⁸, set up a group of experts tasked with developing AI ethics guidelines in a broad alliance of stakeholders from politics, the economy, and science.

European fundamental values will be given due consideration in this development process. This is crucial, and not least so for the current European initiative for the promotion of the development of artificial intelligence technologies of April 8, 2018, which is intended to boost R&D investment in AI by at least 20 billion euros by 2020.⁹ The central issue to be resolved here concerns observance of fundamental EU rights such as fairness, safety, transparency, and solidarity but also moral regulations for the future of labor, and securing the survival of democracy.

Emotional intelligence to control technology

Friedrich Dürrenmatt’s play *The Physicists*, which raises the fundamental question of the responsibility for technological developments that involve a hazard of becoming a threat to the world, thus gains new topicality. Keeping control over technology is a categorical must of all human aspiration—after all, the very essence of being human is about freedom, volition, and creativity.

This entails the question whether present-day education systems can account for those new challenges. Today’s admission criteria for universities and universities of applied sciences currently set the wrong priorities. In many disciplines, extensive factual knowledge is still overly valued, and not enough attention is given to empathy, creativity, and problem-solving competency in incomplete and conflictive situations as well as to holistic views of complex systems.

The point is to revise the education system fundamentally. In order to live with robots and AI in the future and to still stay on top of the technology, it will be necessary to focus on the core human competences of emotional intelligence and true education and to replace trivial fact learning with soft skills. You cannot compete with robots and AI—you have to control them. ✕

⁸ See “Artificial intelligence: Commission kicks off work on marrying cutting-edge technology and ethical standards,” press release of the European Commission, March 9, 2018; http://europa.eu/rapid/press-release_IP-18-1381_en.htm (May 14, 2018).

⁹ See “Artificial intelligence: Commission outlines a European approach to boost investment and set ethical guidelines,” press release of the European Commission, April 25, 2018; www.europa.eu/rapid/press-release_IP-18-3362_en.pdf (May 14, 2018).

Allan Hanbury im Gespräch mit Martin Kugler

»Was man mit Daten tun kann,
hat seine Grenzen«



Allan Hanbury, der im Bereich Data Intelligence arbeitet, erklärt, was man mithilfe künstlicher Intelligenz aus Daten herauslesen kann – und was nicht. Er betont die Wichtigkeit der Technologiefolgenabschätzung für alle, die in der Grundlagenforschung tätig sind.

Sie sind jüngst zum Professor für Data Intelligence an der Technischen Universität Wien ernannt wurden. Was ist Data Intelligence?

Allan Hanbury: Data Intelligence nimmt riesige Mengen in vielen Anwendungsbereichen aktuell verfügbarer Daten und leitet daraus Einsichten ab, die für evidenzbasierte Entscheidungen genutzt werden können.

Zu welchen Zwecken?

AH: Zwecke gibt es viele. Das hängt vom Bereich ab. Wir haben viel zum Thema Patente, zum Thema Rechte an geistigem Eigentum gearbeitet. Die Geschichte der Patente umspannt mehr als hundert Jahre und enthält Informationen über die Technik der Menschheit. Das muss man strukturieren, und zwar so, dass sich dadurch Aufschlüsse über die Entwicklung der Technik ergeben. Diese bilden dann die Grundlage für die Entscheidung, ob ein neues Patent gewährt werden soll oder nicht. Wir haben auch im medizinischen Bereich gearbeitet, in dem regelmäßig Patientendaten gesammelt und irgendwie aufbewahrt werden und Publikationen die Forschung auf diesem Gebiet dokumentieren. Neue Medikamente oder Behandlungen beispielsweise werden im klinischen Versuch erprobt, aber sobald sie zum Einsatz zugelassen sind, werden Wirksamkeit und Nebenwirkungen weiterhin in Krankenakten dokumentiert. So entsteht eine Menge Information, die zurzeit nicht optimal genutzt wird. Die Daten werden an verschiedenen Stellen aufbewahrt und bedürfen der Bearbeitung, bevor man sie analysieren, zusammenbringen und bestmöglich nutzen kann. Auch in Unternehmen gibt es riesige Datenmengen, die heute nicht optimal Verwendung finden. Unterschiedliche Abteilungen eines Unternehmens haben oft eigene Systeme, und die Daten sind voneinander getrennt. Verbindet man

Allan Hanbury ist Professor für Data Intelligence an der Technischen Universität Wien und Mitglied des Lehrkörpers des Complexity Science Hub Vienna. Er leitet das IKT-Leuchtturmprojekt »Data Market Austria«. Er fungierte als wissenschaftlicher Koordinator des von der EU finanzierten Projekts »Khresmoi«, das sich mit der Suche nach Informationen über Medizin und Gesundheit und deren Analyse befasst, und als Mitbegründer des Spin-off-Unternehmens »contextflow«, das die »Kreshmoi«-Technologie vermarktet, mit der Röntgenbilder abgeglichen werden können. Zudem koordinierte er die EU-finanzierten Projekte »VISCERAL« über die Bewertung von Algorithmen im Bereich Big Data und das sich der Analyse medizinischer Texte widmende Projekt »Kconnect«.

Allan Hanbury is Professor for Data Intelligence at Vienna's University of Technology and faculty member of the Complexity Science Hub Vienna. He leads the ICT lighthouse project *Data Market Austria*. He was scientific coordinator of the EU-funded *Khresmoi* project on medical and health information search and analysis and is co-founder of *contextflow*, the spin-off company commercializing the *Khresmoi* radiology image search technology. He also coordinated the EU-funded *VISCERAL* project on evaluation of algorithms on big data and the EU-funded *Kconnect* project on analyzing medical text.

Allan Hanbury in an interview with Martin Kugler

“There are limits to what can be done with data”

Allan Hanbury, researcher in the field of data intelligence, explains what can be read out of data with the help of artificial intelligence—and what is out of reach. He emphasizes the significance of technology assessment for people doing basic research.

You are newly appointed Professor for Data Intelligence at Vienna's University of Technology. What is “data intelligence”?

AH: Data intelligence takes huge amounts of data currently available in many application areas and extracts insights out of them that can be used for evidence-based decisions.

For which purposes?

AH: There are many purposes. It depends on the area. We've done a lot of work on patents, on intellectual property rights. There are over one hundred years of patents that contain information about the technology of humanity. It is very important to structure this in a way to shed light on how this technology developed. This forms the basis for the decision on whether a new patent should be granted or not. We also worked in the medical domain, with data collected regularly from patients and stored in some way and with publications documenting the research in the field. A new medication or treatment is tested in a clinical trial, for example, but once it is certified for use on patients, the efficacy and side effects continue to be documented in patient records. This generates a lot of information that is currently not optimally used. It is stored in many places, and some processing is needed to be able to analyze it, to bring it together, and to use it in the best way. Inside companies, there is also a huge amount of information that is currently not used optimally. Different departments of a company often have their own systems; the data are separated from each other. By bringing data together, it is possible to get useful insights. We have done a pro-

sie, lassen sich nützliche Einsichten gewinnen. Wir haben etwa ein Projekt im Bereich Industrie 4.0 durchgeführt, in dessen Rahmen wir von Mitarbeitern eingegebene Informationen über Reparaturen von Fehlern an Maschinen analysiert haben, um weniger erfahrenen Mitarbeitern Lösungen anbieten zu können.

Gehen wir einen Schritt zurück: Was genau bedeutet der Begriff »Intelligenz« in diesen Zusammenhängen?

AH: Große Mengen von Informationen sammeln und analysieren und daraus nützliche Erkenntnisse gewinnen.

Welche Rolle spielt künstliche Intelligenz (KI) bei diesem Verfahren? Bedeutet sie dasselbe wie »Datenintelligenz«.

AH: KI ist der Intelligenz eines menschlichen Gehirns ähnlicher. Sie ist eine Rekonstruktion der Intelligenz des menschlichen Gehirns in Form eines künstlichen Systems. Das Problem bei KI ist, dass wir es hier mit einem sehr weiten Begriff zu tun haben; das Etikett KI umfasst zahlreiche analytische Zugänge. Unglücklicherweise hat KI auch eine sehr zerrissene Geschichte. Sie nahm in den 1950er-Jahren ihren Anfang, und im Lauf der Zeit wurden viele Versprechungen gegeben – Versprechungen, die allerdings nicht erfüllt wurden. Dann kam es mehrere Jahrzehnte hindurch zu einer Reihe von Wintern in diesem Bereich; Förderungen wurden zurückgenommen, und nur wenige fanden den Mut, sich auf diesem Gebiet zu engagieren. Vor Kurzem hat KI wieder abgehoben. Auf einmal ist KI ein trendiger Begriff, was nicht ganz unproblematisch ist. Seth Juarez hat gesagt: »Wir setzen auf KI, wenn wir etwas verkaufen wollen. Wir reden über Maschinenlernen, wenn wir jemanden anstellen wollen.« KI wird eher als Marketingbegriff verwendet.

Sie arbeiten im Bereich Maschinenlernen?

AH: Ich arbeite an der Schnittstelle zwischen Theorie und Praxis. In theoretischer Hinsicht können die verschiedensten Ansätze für unterschiedliche Dinge verwendet werden; damit meine ich Teilgebiete der KI wie Textanalyse, Maschinenlernen und Informationsgewinnung. Andererseits sehen wir uns in einigen Anwendungsbereichen wie etwa der Medizin oder der Industrie mit ganz spezifischen Problemen konfrontiert. Es ist nicht wirklich so einfach, dass man sich einen Algorithmus aus dem Regal holt und diesen auf ein Problem aus dem wirklichen Leben anwendet. Zuerst gilt es das Problem so zu definieren, dass es sich lösen lässt, was sich schwierig gestalten kann, weil es in vielen Bereichen implizite Informationen in Form eines gemeinsamen Ver-

ject in the area of Industry 4.0, for instance, in whose context we analyzed information entered by employees about how they repaired faults in machines to be able to suggest solutions to less experienced employees.

Taking one step back: What does the term “intelligence” exactly mean in these contexts?

AH: It means to gather and analyze a large amount of information and get some kind of useful insights out of it.

What is the role of artificial intelligence (AI) in this process? Does it mean the same as “data intelligence”?

AH: AI is more like the intelligence of the human brain. It rebuilds the human brain intelligence in an artificial system. The problem of AI is that it is a very broad term; the label AI includes many analytic approaches. Unfortunately, AI also has a patchy history. It began in the 1950s, and many promises have been made over time. These promises were not really followed through. Then there was a series of AI winters for several decades when research funding was cut back and only few people were brave enough to work in the field. Then AI took off again recently. Suddenly it is an in-word, and this is a bit of a problem. Seth Juarez said: “We do AI when we want to sell something. We talk about machine learning when we want to hire somebody.” AI is often used as more of a marketing term.

So you do machine learning?

AH: I’m working at the intersection between theory and practice. On the theoretical side, huge numbers of approaches can be used for various things; these include subparts of AI like text analysis, machine learning, and information extraction. On the other side, there are very specific problems in a number of application domains, for example medicine or industry. It is actually not that straightforward: you don’t just take an algorithm off the shelf and apply it to a real-life problem. One first has to formulate the problem in a way that it can be solved—which is made difficult as many domains have implicit information as a common understanding, a specialized terminology, as well as shared assumptions. If you think about medicine, there are certain accepted basics about the human body that everybody who works in this field understands and never explicitly writes down. Knowing these is important for understanding and analyzing the data in a sensible way.

ständnisses, einer speziellen Terminologie oder geteilter Annahmen gibt. Auf dem Gebiet der Medizin beispielsweise gibt es bestimmte anerkannte Grundeinsichten den menschlichen Körper betreffend, die jeder, der in diesem Bereich arbeitet, versteht und nie ausdrücklich festhält. Dieses Grundwissen zu kennen ist wichtig, will man die Daten verstehen und vernünftig analysieren.

Was lässt sich aus Daten herauslesen?

AH: Man braucht nur in die Zeitung zu schauen, um dahinterzukommen, was sich Daten entnehmen lässt. Denken wir etwa an den Facebook-Skandal. Da stellt sich einerseits das Problem der Privatsphäre. Die Europäische Kommission setzt hier gute Schritte in Richtung eines neuen Datenschutzgesetzes – was sehr begrüßenswert ist. Andererseits gibt es Leute, die riesige Mengen persönlicher Informationen online stellen. Das macht es bestimmten Unternehmen leicht. Stehen einem die Daten von Facebook zur Verfügung, kann man sich ein sehr detailliertes Bild von jedem seiner Kunden machen. Obwohl viele Firmen vielleicht nicht über das Maß an Information von Facebook verfügen, können sie sich auf Grundlage der ihnen verfügbaren Daten einen guten Begriff vom Verhalten ihrer Kunden machen. Man denke an die Treuekarten von Supermärkten, die genau wissen, was deren Inhaber kaufen. Sind genügend Daten verfügbar, kann man verstehen, um welche Leute es sich handelt und was sie tun – und sogar mit welchen Persönlichkeiten man es zu tun hat. Es gibt für das Wohl der Menschheit nützlichere Einsichten, etwa im Bereich der Medizin. Die Analyse von Patientenakten hat unerwartete Nebenwirkungen von Medikamenten aufgezeigt. Jedes Medikament muss eine Testphase durchlaufen, aber vielleicht zeitigt die Verabreichung an ein paar hundert Personen in keinem Fall eine Nebenwirkung. Kommt das Medikament dann auf den Markt, wird es bei Hunderttausenden Menschen eingesetzt: Dann lassen sich möglicherweise Nebenwirkungen ersehen, die sich durch eine Analyse der Patientenakten nicht haben entdecken lassen. Ich kenne mindestens einen Fall, in dem ein Medikament aufgrund einer aus medizinischen Aufzeichnungen gewonnenen Einsicht vom Markt genommen wurde. Solche Beispiele könnte ich noch viele nennen.

What can be read out of data?

AH: You just have to read the newspapers to work out what people can read out of data. Take the latest Facebook scandal as an example: On the one hand, you have the privacy issue. There are very good steps made by the European Commission for the new data protection legislation—which is very welcome. On the other hand, there are people who put huge amounts of personal information online. This makes it easy for certain companies. If you have the data that Facebook has, you can get an extremely detailed understanding of each of your customers. Though many companies may not have the same level of information that Facebook has, they can get a good picture of how customers behave based on the data they have. Think of supermarket loyalty cards; supermarkets know exactly what you buy. With enough data, you can reach some understanding of what people are and what they do—or even of their personalities. There are insights that are more useful for the good of humankind, for example in medicine. Analyses of patient records have disclosed unexpected side effects of medications. Every medication has to go through a trial, but this trial could be on a few hundred people, and maybe none of them shows a side effect. Once the drug gets rolled out, it is used on hundreds of thousands of people, then there are potentially more cases of side effects that can be detected by analyzing the patient records. I know at least one example where a drug has been taken off the market because of such an insight from medical records. I could keep on giving you more examples.

All this information you get out of data depends on their quality. How can you guarantee the accuracy of huge amounts of data?

AH: It depends on the application. Consider recommender systems in online shops: The companies know what you bought and make recommendations. No company is sued for bad recommendations. It doesn't matter if the companies are right. They do their best and if they're lucky the customer will buy something that the company recommended. And if not, nobody will complain. It would be very important that the data is of very high quality in certain areas, but it isn't—for example in the area of medicine, where much of the information is stored in unstructured form and the quality is not optimal for getting the insights needed. There are ways to test this, you can evaluate the material by doing experiments with algorithms and data to see if it produces

Die gesamte aus Daten gewonnene Information hängt von der Qualität der Daten ab. Wie kann man die Genauigkeit riesiger Datenmengen garantieren?

AH: Das kommt auf die Anwendung an. Denken Sie an Empfehlungssysteme in Onlineshops: Die Unternehmen wissen, was man gekauft hat, und empfehlen einem bestimmte Produkte. Keine Firma wird wegen einer schlechten Empfehlung geklagt werden. Es macht nichts, wenn sie nicht richtigliegt. Sie tut ihr Bestes, und wenn sie Glück hat, kauft der Kunde etwas, was sie empfohlen hat. Wenn nicht, wird sich niemand beklagen. Entscheidend wäre es, dass die Daten in bestimmten Bereichen sehr gut sind, was aber nicht der Fall ist. Im medizinischen Bereich etwa werden viele Informationen in unstrukturierter Form gespeichert; die Qualität der Daten ist nicht optimal, wenn es darum geht, sich die erforderlichen Einsichten zu verschaffen. Man kann das testen, kann das Material mit Algorithmen evaluieren, um festzustellen, ob die Daten ergeben, was sie ergeben sollen. Wo ein Algorithmus auf denkbare Ausreißer verweist, die man sich dann genauer anschauen kann, gibt es Möglichkeiten, mangelhafte Datensätze zu erkennen. In diesem Fall ist es gut, auf menschliche Expertise zurückzugreifen, wo das notwendig ist. Sogar von Menschen gekennzeichnete Daten sind nicht immer genau. Ein Beispiel: Ich habe mit Radiologen an einem Projekt gearbeitet. Sie bekamen Computertomografiebilder, die den Rumpf von Patienten zeigten, und sollten die verschiedenen Organe genau umreißen. Es waren drei Radiologen, welche sich unabhängig voneinander mit den Bildern auseinandersetzten. Es ergaben sich tatsächlich gravierende Unterschiede in der Auffassung dessen, was noch zu einem Organ gehört und was nicht. Das mag ein ganz spezifisches Problem sein und nicht ins Gewicht fallen, wenn es darum geht, ob ein Leberschaden vorliegt oder nicht, für eine Algorithmen lernende Maschine allerdings ist es ganz entscheidend, hier Klarheit zu haben.

In Zukunft wird es viel mehr Daten geben: von Verkaufssystemen, Sensornetzwerken, selbstfahrenden Autos usw. Sind der Aneignung von Daten Grenzen gesetzt?

AH: Ich sehe drei Grenzen. Erstens die Privatsphäre: Was man mit Daten tun kann, hat seine Grenzen. Die zweite Grenze ist technischer Natur: Hier geht es darum, Daten zu verbinden, damit man sie gemeinsam analysieren kann. Die dritte Grenze betrifft die Menschen, in deren Besitz die Daten sind: Wollen sie die Daten tatsächlich auf sinnvolle Weise verbunden sehen? Hier kommt es auf den politischen Willen und wirtschaftliche Interessen an.

what it should produce. There are ways of detecting poor data where an algorithm can flag potential outliers in the data and somebody can look at it. A good approach is getting human expertise involved where necessary. Even data labelled by people is not always accurate. To give an example: I was involved in a project together with radiologists. We gave them computer tomography images of the torso and they were asked to accurately demarcate the different organs. We asked three radiologists independently to analyze the images. There were some major differences between what they considered to be part of each organ and what not. This is a very specific problem and maybe not so important for being able to decide if the liver is damaged, but for machine learning algorithms this is a key thing to know.

There will be much more data in the future: from retail systems, from sensor networks, from autonomous cars, etc. Are there any limits to the acquisition of data?

AH: I see three limits. One is privacy: there are limits to what can be done with data. The second limit is a technical one: what is actually possible in terms of bringing these data together in a way they can be analyzed together. The third limit is the will of the people in possession of the data to see them actually brought together in a way that makes sense. This could be political will or economic interests.

Where are the problems concerning the technical issues?

AH: The main problems are with legacy data: data that already exist and have to be brought together. Solving this is often not easy. If you think of designing new systems, you always need to create some sort of data structure. Usually this data structure is designed for a specific system without thinking of long-term implications concerning reuse of the data. Sometimes it is impossible to match records from different systems.

But there should be solutions to that kind of problem...

AH: Yes, usually there are solutions, but it takes a while.

Wo liegen hier die technischen Schwierigkeiten?

AH: Das Hauptproblem ist das Datenerbe: Daten, die es bereits gibt und die zusammengeführt werden müssen. Hier eine Lösung zu finden ist oft schwierig. Plant man ein neues System, muss man immer eine bestimmte Datenstruktur schaffen. Im Allgemeinen ist eine solche Datenstruktur für ein bestimmtes System ausgelegt und nimmt auf keine langfristigen Implikationen wie die Wiederverwendung von Daten Bedacht. Manchmal ist es schlichtweg unmöglich, Aufzeichnungen aus verschiedenen Systemen aufeinander abzustimmen.

Für solche Probleme sollte es aber eine Lösung geben...

AH: Die gibt es auch in der Regel. Aber es dauert einige Zeit, solche Lösungen zu erarbeiten.

Diese Methoden der Sammlung und Analyse von Daten flößen vielen Menschen Furcht ein. Ist ihre Angst berechtigt?

AH: Es gibt verschiedene Arten der Angst. Erstens fürchten Menschen, dass ihre Daten missbraucht werden können. Zweitens haben Menschen Angst davor, dass ihre Jobs wegautomatisiert werden. Und drittens kann man sich auch vor Systemen fürchten, die wie Skynet in den *Terminator*-Filmen die Menschheit zerstören. Vor Skynet, finde ich, braucht man keine Angst zu haben. Angesichts heutiger KI-Ansätze wird Skynet wohl nicht so bald auf uns zukommen. Nicht so einfach ist es im Hinblick darauf, was gelernt werden kann. Bei Algorithmen für Maschinenlernen bedarf es ganz bestimmter Inputs und Outputs sowie einer Person, die das System trainiert. Das System tut einfach, was man ihm beigebracht hat.

Und wie steht es um den Missbrauch von Daten?

AH: Davor Angst zu haben ist berechtigt. Hier gibt es zwei Gegenmaßnahmen. Einerseits eine entsprechende rechtliche Regelung, wie sie gerade im Entstehen ist. Das ist aber nur eine Seite. Die andere Seite ist, dass Menschen für ihre Daten Verantwortung übernehmen und sich dessen bewusst werden, wer wozu Zugang hat und was womit geschieht. Das könnte man an Schulen unterrichten, was aber zurzeit nicht geschieht. Viele Menschen sind sich des Werts ihrer Daten nicht bewusst. Es gab ein paar Initiativen, die es einzelnen Personen erlaubten, Kontrolle über ihre persönlichen Daten auszuüben und sie zu verkaufen, aber das ging schief. Es gibt kein Bewusstsein. Noch immer überlassen die Leute Facebook, Google und anderen Diensten freizügig ihre Daten.

Many people are scared of all these methods to collect and analyze data. Are they right to be afraid?

AH: They may be afraid in multiple ways. The first is that data is misused. The second is that their jobs will be automated away. And actually there is a third one: being afraid of a system like Skynet in the *Terminator* movies that destroys humanity. I would say that this third field is nothing to be afraid of. Given the state of today's AI approaches, we are not going to see Skynet so soon. It's not so simple in terms of what can be learned. If you think of machine learning algorithms, you have to have very specific inputs and very specific outputs and somebody to train the system. It just does what it has been trained to do.

What about the misuse of data?

AH: The misuse of data is a thing to be afraid of. There are two ways to work against that. An adequate legislation, which is on its way, is just one side. The other is that people take responsibility for their data and are aware of who has access to what and what is done with it. This is actually something that could be taught at schools—which is currently not the case. Many people are not aware of the value of their data. There have been a few initiatives allowing individuals to control and sell their personal data, which unfortunately have failed. There is no awareness. People still give their data free to Facebook, Google, and other services.

And the third concern: that jobs are automated away. Is this a major issue?

AH: Yes, it is. The profiles of available jobs change. This has happened ever since the Industrial Revolution. When I recently took a taxi from the airport, the driver asked me what I do for a living. So I told him. He probed me for an estimate of when self-driving cars are going to be on the road. I gave him a number. He replied, "At least I will be in pension by then." Job profiles are changing faster in some areas than in others. I think that the knowledge work areas will take longer. Their systems will initially support us in being more efficient and not replace us. Maybe in less skilled jobs it will happen more rapidly. Even self-driving cars are not coming as rapidly as has been predicted. Some time ago, the time horizon was five years. But there are still problems.

Und wie steht es um die dritte Sorge, dass Jobs wegautomatisiert werden? Ist das ein Problem, das ins Gewicht fällt?

AH: Ja. Das Profil der angebotenen Jobs ändert sich. Das ist seit der industriellen Revolution so. Als ich vor Kurzem am Flughafen ein Taxi nahm, fragte mich der Fahrer nach meinem Beruf. Ich erzählte ihm, was ich mache. Er wollte wissen, wann ungefähr selbstfahrende Autos auf den Straßen unterwegs sein werden. Ich sagte ihm, wann das meiner Schätzung nach der Fall sein wird. Er antwortete: »Da bin ich dann wenigstens schon in Pension.« Die Jobprofile ändern sich in manchen Bereichen schneller als in anderen. Ich glaube, dass das in den Bereichen der Wissensarbeit länger dauern wird. Dort werden uns die Systeme anfangs dabei unterstützen, effizienter zu sein, und uns nicht ersetzen. In weniger qualifizierten Berufen wird es schneller gehen. Aber nicht einmal bei selbstfahrenden Autos geht es so schnell wie vorhergesagt. Vor einiger Zeit war der Zeithorizont fünf Jahre. Doch es gibt noch immer Probleme.

Sollten sich Wissenschaftler wegen der Nebenwirkungen von KI Sorgen machen?

AH: Ganz wichtig ist, dass sich alle, die in der Grundlagenforschung tätig sind, dessen bewusst sind, was Technologiefolgenabschätzung bedeutet.

Bedarf die Wissenschaft der Führung, um diesen Gefahren zu entgehen?

AH: Die Wissenschaft hat Führung, und zwar nicht dadurch, dass man Wissenschaftlern erlaubt, bestimmte Dinge zu tun, sondern durch Anreize, die für bestimmte Themen Förderungen zugänglich machen. Regierungen und EU-Behörden stellen gewisse Mittel für Arbeiten zu diesem oder jenem Thema zur Verfügung, und das ist gut so. Es ist allerdings immer empfehlenswert, Wissenschaftler an Dingen arbeiten zu lassen, die keine unmittelbare Anwendung haben, aber sich vielleicht später umsetzen lassen.

Haben Sie eine Vorstellung, wie die Zukunft aussehen wird? Wie werden Menschen und Maschinen zusammenarbeiten und miteinander leben?

AH: Meine kurzfristige Vision ist die einer Welt, in der Computer Wissensarbeiter bei verschiedenen Aufgaben unterstützen. Wenn man heute etwas wissen will, muss man eine Suchmaschine aufrufen, seine Suchkette eintippen und so weiter. Das nimmt nicht nur Zeit in Anspruch, sondern lenkt einen auch von den zu lösenden Aufgaben ab. Meine Vision wäre die, dass ein Computerprogramm fortlaufend

Should scientists be concerned about the side effects of AI?

AH: It is very important that people doing basic research are aware of technology assessment.

Should science be guided to avoid some of the threats?

AH: Science is actually guided, not by not allowing scientists to do things, but through incentives making funding accessible for specific topics. Governments and European Commission authorities make a certain amount of money available to work on this or that. This is fine. However, it is always good to have scientists working on things that have no immediate application but may later turn out to have one.

Do you have a vision of the future? How will human beings and machines collaborate and live together?

AH: My short-term vision is a world in which computers assist knowledge workers in various tasks. If you want to know something today, you have to go to a search engine, type your search string in, and so on. This takes time and actually distracts you from the tasks that you have to solve. The vision here is that a computer program continuously observes what the user does and provides the information they need at the right time. This should happen in a way that doesn't distract, but is always there. I am involved in a spin-off-company that is implementing exactly this for radiologists: The radiologist can select an area of the image and is immediately provided with information on potential pathologies associated with the selected pattern based on an analysis of stored images. The next step would be that the system suggests to the radiologist what is unusual and returns the necessary information supporting this. This vision of how knowledge workers and computers may work together can be achieved within the next years.

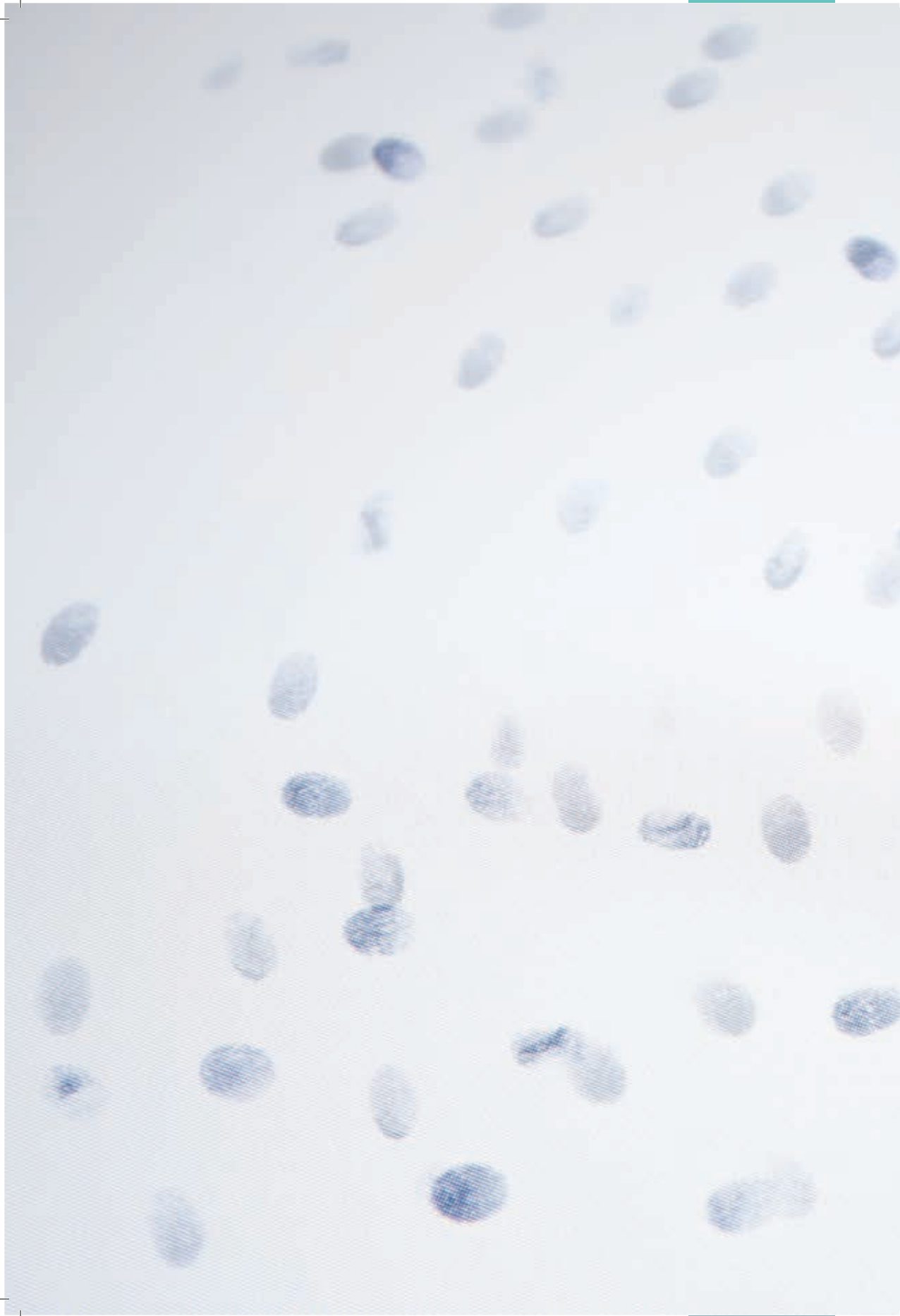
beobachtet, was der Nutzer tut, und diesem zum richtigen Zeitpunkt die Information zur Verfügung stellt, die er braucht. Das sollte auf eine Art geschehen, die nicht ablenkt, aber ständig präsent ist. Ich engagiere mich in einer Ausgründung, die gerade genau das für Radiologen umsetzt: Der Radiologe wählt einen bestimmten Bildausschnitt und erhält unmittelbar Informationen über möglicherweise pathologische Gegebenheiten, die auf Grundlage einer Analyse der gespeicherten Bilder mit dem gewählten Muster in Zusammenhang stehen. Der nächste Schritt wäre dann, dass das System den Radiologen darauf hinweist, was ungewöhnlich ist, und die notwendigen Informationen liefert, die das belegen. Diese Vision der Zusammenarbeit zwischen Wissensarbeitern und Computern kann in den nächsten Jahren Wirklichkeit werden.

Und längerfristig?

AH: Das ist schwer zu sagen. Denken Sie an breitere Anwendungen wie alte oder kranke Menschen unterstützende Roboter. Das war immer eine der großen KI-Visionen. Alle Anwendungen, die ich gesehen habe, sind von völlig autonomen Computern, die alles verstehen und tun, weit entfernt. Roboter oder KI-Systeme führen nur die Aufgaben aus, für die sie gebaut wurden. Im Lauf der Zeit wird die Bandbreite der durchführbaren Aufgaben größer werden. Es wird jedoch immer bestimmte Gesichtspunkte und Bereiche geben, bei denen nicht ganz klar ist, was vom System erwartet wird und wie man feststellen soll, wie weit es den Anforderungen entspricht. ✕

And for the long term?

AH: This is hard to say. Look at broader applications like robots assisting old or sick people—which has always been one of the big visions of AI. All the applications I have seen are quite simple, they are far from complete autonomous robots that understand and do everything. A robot or an AI system performs only the specific tasks for which it was built. Over time, the tasks that can be carried out will expand. Yet there will always be some areas and some aspects where it is not quite clear what is expected from the system or how to assess the degree to which a system meets the requirements. ✕



**P
E
T:**

»l
b
B
e
s
is
ü
F
B
M
n
g
a
s

“P
C
d
fi
T
P
P
a
r
o
w
w

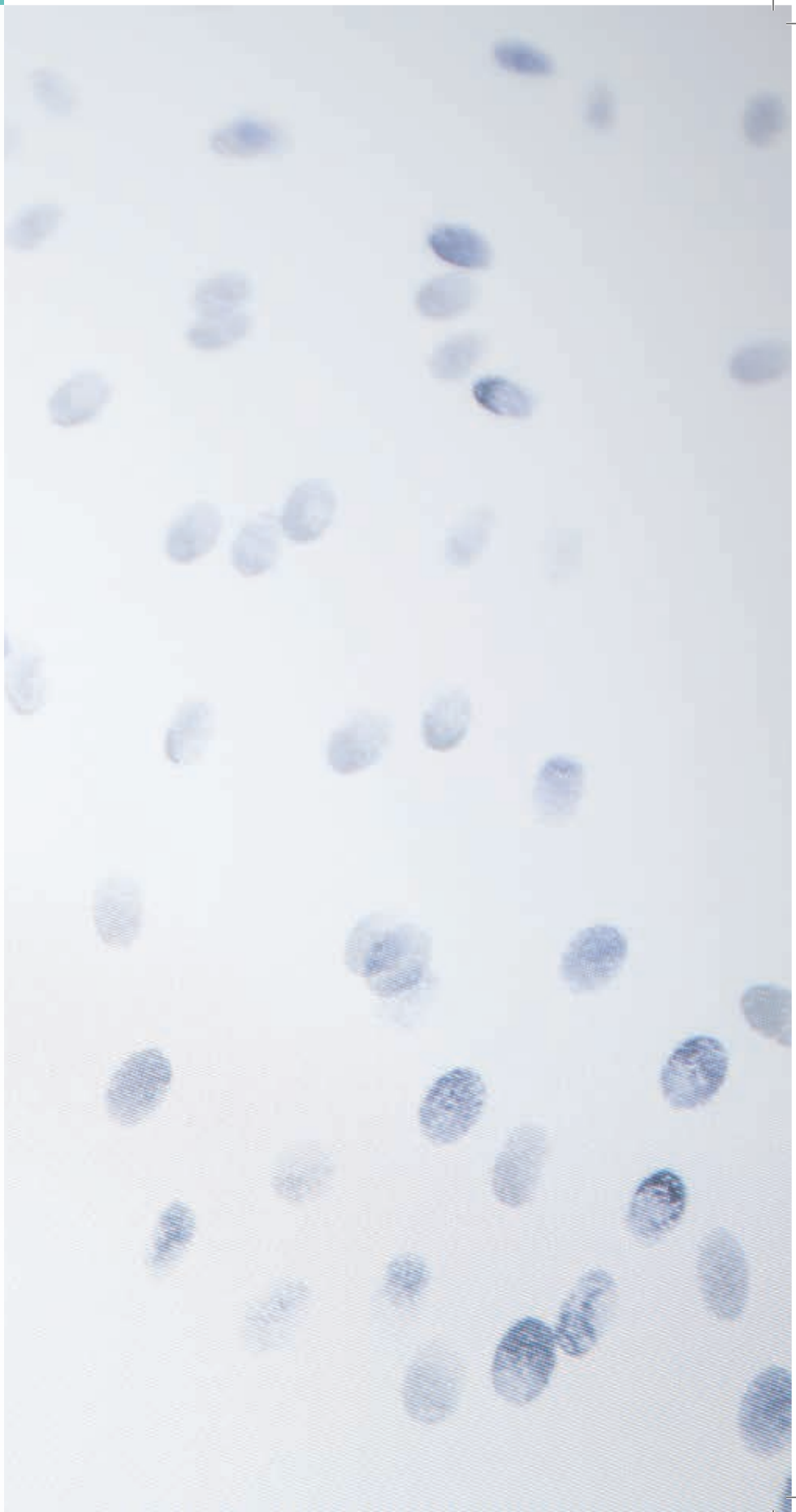
—
©
t

Pool of Fingerprints
Euclid/Masahiko Sato &
Takashi Kiriya (JP)

»Pool of Fingerprints« besteht aus einer großen Bildschirmfläche und einem Fingerabdruckscanner. Die Oberfläche ist mit Fingerabdrücken übersät, die wie ein Fischschwarm über den Bildschirm schwimmen. Man kann seine eigenen Fingerabdrücke abgeben und mit den anderen schwimmen sehen.

“Pool of Fingerprints” consists of a large display surface and a fingerprint scanner. The display surface is populated with fingerprints swimming like a school of fish. Visitors may surrender their own fingerprints and watch them swim with others.

© Ars Electronica/
tom mesic



Die Arbeit der Zukunft oder Die Grenzen des Maschinenlernens

Künstliche Intelligenz wird einen Teil der bisher als spezifisch menschlich erachteten Fähigkeiten übernehmen – und damit auch unsere Arbeitswelt verändern. Beim derzeitigen Entwicklungsstand maschineller Lernalgorithmen sind dem aber Grenzen gesetzt. Die beiden us-Forscher Erik Brynjolfsson und Tom Mitchell haben acht Kriterien definiert, in welchen Fällen Methoden der künstlichen Intelligenz sinnvoll einsetzbar sind.

Computer haben in den vergangenen Jahrzehnten die Arbeitswelt grundlegend verändert – und zwar in praktisch jedem Wirtschaftssektor: Es gibt mittlerweile kaum mehr einen Beruf, in dem keine digitalen Technologien eingesetzt werden. Das war aber erst der Anfang: »Aufgrund der jüngsten Fortschritte des maschinellen Lernens, das fähig ist, die Automatisierung selbst zu beschleunigen, stehen wir nun am Beginn einer noch größeren und schnelleren Transformation«, sind die beiden us-Forscher Erik Brynjolfsson (Massachusetts Institute of Technology/MIT) und Tom Mitchell (Carnegie Mellon University) überzeugt. Wie sie in einem im Dezember 2017 in *Science* erschienenen und weithin beachteten Artikel betonen, sei allerdings derzeit nicht klar, bei welchen Aufgaben sich das maschinelle Lernen wirklich auszeichnen könne, weshalb es auch nur wenig Einigkeit über die spezifisch erwarteten Auswirkungen auf die Arbeitskräfte und auf die Wirtschaft gebe.

Die beiden Forscher haben sich in dem Artikel das Ziel gesetzt, einige allgemeingültige Aussagen zu den Auswirkungen von Maschinenlernen zu treffen – und zwar basierend auf heutigen Technologien und Erfahrungen. Bevor sie ins Detail gehen und untersuchen, welche Jobs bzw. Teile von Jobs für Maschinenlernen »geeignet« sind, halten sie zwei grundlegende Punkte fest, die zur Zeit wohl unbestritten sind. Erstens: »Wir sind sehr weit von einer künstlichen allgemeinen Intelligenz entfernt. Maschinen sind nicht imstande, die gesamte Palette von Aufgaben zu erledigen, die von Menschen erfüllt werden können.« Das hänge eng mit der zweiten unbestrittenen Tatsache zusammen – nämlich mit dem sogenannten »impliziten Wissen«, das wir Menschen haben. »Wir wissen mehr, als wir sagen können.« So seien etwa das Erkennen eines Gesichtes, Radfahren oder das Verstehen von gesprochener Sprache Aufgaben, die der Mensch sehr gut kann. Wir könnten jedoch nur schwer reflektieren, wie wir diese Aufgaben ausführen. »Wir können viele Aufgaben nicht leicht, wenn überhaupt, in formalen Regeln kodifizieren.« Durch fortgeschrittene Technologien könnte sich das zwar in Zukunft ändern, aber dennoch bleiben für die beiden Forscher Zweifel, ob alle Lernaufgaben ausreichend genau definiert werden können.

The Work of the Future or The Limits of Machine Learning

Artificial intelligence will ultimately perform parts of the skills that have previously been considered specifically human—and consequently also transform our working world. However, given the current development of machine-based learning algorithms, this has its limits. The two us scientists Erik Brynjolfsson and Tom Mitchell have defined eight criteria as to a sensible application of the methods of artificial intelligence.

In the past decades, computers have profoundly changed the employment situation in almost every economic sector: meanwhile there is hardly a business or profession that does not rely on digital technologies. But this was only the beginning: “We are now facing an even larger and more rapid transformation due to recent advances in machine learning, which is capable of accelerating the pace of automation itself,” Erik Brynjolfsson (Massachusetts Institute of Technology/MIT) and Tom Mitchell (Carnegie Mellon University) are convinced. As they emphasize in a widely discussed article published in *Science* in December 2017, it is not yet entirely clear in what tasks machine learning can actually excel, so that there is also little agreement as yet on the specific impacts it can be expected to have on the workforce and economy.

In their article, the two scientists make a number of general statements on the consequences of machine learning—that is, starting out from current technologies and experiences. Before going into detail and looking into what jobs or parts of jobs might be “suitable” for machine learning, they recognize two fundamental considerations that seem to find general acceptance at present: “We remain very far from artificial general intelligence. Machines cannot do the full range of tasks that humans can do.” This is evidently closely linked to a second uncontested fact—namely that of the so-called “implicit knowledge” humans possess: “We know more than we can tell.” For example, recognizing a face, riding a bike, and understanding speech are tasks humans can do extremely well. But, as the two authors point out, it is difficult for us to reflect on how we perform these tasks. “We cannot codify many tasks easily, or perhaps at all, into a set of formal rules.” This could probably change in the future thanks to the application of advanced technologies, but for the two scientists it remains doubtful whether it will indeed be possible to accurately define the whole range of tasks to be learned.

Acht Kriterien für effektives Maschinenlernen

Welche Aufgaben sind nun – aus heutiger Sicht und mit heutigen Methoden – für maschinelles Lernen geeignet? Brynjolfsson und Mitchell haben acht Schlüsselkriterien für einen erfolgreichen Einsatz von Lernalgorithmen herausgearbeitet.

1. **Die Lernfunktion bildet gut definierte Inputs in gut definierten Outputs ab.** Zu den erfolgreichen Beispielen für Maschinenlernen gehören unter anderem die Klassifizierung (z. B. die Zuordnung von Bildern von Hunderassen oder die Kennzeichnung von Krankenakten gemäß der Wahrscheinlichkeit für eine Krebserkrankung) oder Vorhersagen (z. B. die Analyse eines Kreditantrags auf die Wahrscheinlichkeit eines zukünftigen Ausfalls hin). Dabei bleibe es aber stets bei gelernten statistischen Korrelationen; kausale Effekte könnten nicht erfasst werden, so die beiden Forscher.
2. **Es gibt große (digitale) Datensätze, die Input-Output-Paare enthalten.** Je mehr Trainingsbeispiele verfügbar sind, desto fehlerfreier ist das Lernen. Besonders wichtig ist, dass alle relevanten Eigenschaften objektiv in den Trainingsdaten erfasst sind, denn Computer seien anfällig für Nachahmung und unerwünschte Voreingenommenheiten in den Trainingsdaten.
3. **Die Aufgabe liefert eine klare Rückmeldung mit klar definierbaren Zielen und Metriken.** Maschinenlernen funktioniert gut, wenn wir die Ziele klar beschreiben können – auch dann, wenn wir nicht angeben können, wie diese Ziele am besten zu erreichen sind. Dies stehe im Gegensatz zu früheren Ansätzen der Automatisierung, betonen Brynjolfsson und Mitchell. Die Fähigkeit, Input-Output-Entscheidungen von Menschen zu erfassen, führt nicht zwangsläufig zu einer optimalen systemweiten Performance, weil Menschen ihrerseits nur unvollkommene Entscheidungen treffen würden. Daher müssten die Trainingsdaten die übergeordneten Ziele beinhalten.
4. **Es darf keine langen Argumentationsketten geben, bei denen Hintergrundwissen oder gesunder Menschenverstand eine Rolle spielen.** Maschinenlernsysteme sind sehr gut beim Lernen von empirischen Zusammenhängen; sie sind aber weniger effektiv, wenn eine Aufgabe lange Argumentationsketten oder komplexe Planung erfordert, die auf Hintergrundwissen oder gesundem Menschenverstand beruhen. So können Maschinen beispielsweise Spiele wie Go oder Schach leicht erlernen, weil diese Aufgaben rasch mit perfekter Genauigkeit simuliert werden können, sodass Millionen von Trainingsbeispielen automatisch gesammelt werden. In den meisten realen Bereichen gibt es aber keine derart perfekten Simulationen.

Eight criteria for effective machine learning

What tasks are then—from today’s perspective and with today’s methods—suitable for machine learning? Brynjolfsson and Mitchell have identified eight key criteria for a successful application of learning algorithms.

1. **The learning function maps well-defined inputs to well-defined outputs.** Successful examples include, among others, classification (e.g., identifying images of dog breeds or labeling medical records according to the likelihood of cancer) and prediction (e.g., analyzing loan applications as to the probability of future default). As the two researchers point out, this is always a learned statistical correlation, which means that causal effects cannot be taken into account.
2. **Large (digital) data sets exist containing input-output pairs.** The more training examples are available, the more error-free the learning will be. It is crucial that all of the relevant features have been recorded objectively, as computers tend to mimic and perpetuate unwanted biases contained in the training data.
3. **The task provides clear feedback with clearly definable goals and metrics.** Machine learning works well if the goals are clearly described—even if we cannot explicitly define the best way of how to achieve these goals. Brynjolfsson and Mitchell emphasize that this contrasts with earlier approaches to automation. The ability to recognize human-made input-output decisions might not necessarily lead to an optimal performance of the whole system, as humans themselves might make imperfect decisions.
4. **Long chains of reasoning involving background knowledge or common sense must be avoided.** Machine learning systems are extremely efficient at learning empirical interdependencies but are less effective if a task requires long chains of reasoning or complex planning based on background knowledge or common sense. For example, machines will quickly learn such games as go or chess because these tasks can easily be simulated with perfect accuracy, so that millions of training examples can be collected automatically. However, in most real-world contexts there are no such perfect simulations.
5. **There is no need for explaining in detail how a decision was made.** Large neural nets learn how to make decisions by subtly adjusting up to hundreds of millions of numerical weights interconnecting their artificial neurons. But it is difficult to follow their reasoning, as machines, unlike humans, rarely make use of the same intermediate abstractions. For example, whereas computers can diagnose cer-

5. **Es besteht keine Notwendigkeit für eine detaillierte Erklärung der getroffenen Entscheidung.** Große neuronale Netze lernen das Treffen von Entscheidungen durch subtile Anpassung von bis zu Hunderten von Millionen von numerischen Gewichten, welche die künstlichen Neuronen verbinden. Die Entscheidung ist aber kaum nachvollziehbar, weil Maschinen häufig nicht die gleichen Zwischenabstraktionen vornehmen wie wir Menschen. Ein Beispiel: Während Computer bestimmte Arten von Krebs oder Lungenentzündung so gut oder sogar besser als Fachärzte diagnostizieren können, ist ihre Fähigkeit zu erklären, wie sie zu dieser Diagnose kamen, sehr schlecht ausgeprägt. Allerdings merken Brynjolfsson und Mitchell an, dass auch Menschen das Lösen mancher Aufgaben nur schwer erklären können – etwa wie wir Wörter in den Geräuschen, die wir hören, erkennen.
6. **Fehlertoleranz und keine Notwendigkeit für nachweislich richtige oder optimale Lösungen.** Fast alle maschinellen Lernsysteme leiten ihre Lösungen statistisch ab. Daher ist es selten möglich, sie zu 100 Prozent genau zu trainieren. Selbst die besten Systeme zur Sprach- und Objekterkennung oder klinischen Diagnose machen Fehler. Daher ist eine Fehlertoleranz der Systeme ein wichtiges Kriterium für den Einsatz von Lernalgorithmen.
7. **Das Phänomen oder die Funktion, die gelernt werden, sollten sich im Laufe der Zeit nicht schnell ändern.** Im Allgemeinen funktionieren maschinelle Lernalgorithmen gut, wenn die Anwendungen den Trainingsbeispielen ähnlich sind. Wenn sich diese im Laufe der Zeit ändern, ist ein erneutes Training erforderlich. Der Erfolg hängt daher von der Änderungsrate ab – und zwar relativ zur Rate des Erwerbs neuer Trainingsdaten. Als positives Beispiel führen die Forscher Spamfilter bei E-Mails an: Die Zahl der unerwünschten Mails ist hoch im Vergleich dazu, wie rasch sich die Art der Spammails verändert.
8. **Es sind keine speziellen Fertigkeiten, körperliche Fähigkeiten oder Mobilität erforderlich.** Roboter sind im Vergleich zu Menschen immer noch ziemlich ungeschickt, wenn sie es mit unstrukturierten Umgebungen und Aufgaben zu tun haben. Das sei allerdings weniger ein Manko des Maschinenlernens als vielmehr eine Folge des Standes der Robotertechnik.

Nur manche Teile eines Jobs sind betroffen

Auf Basis dieser Analyse treffen Brynjolfsson und Mitchell einige allgemeine Aussagen über die Rolle von künstlicher Intelligenz im Arbeitsleben der Zukunft. »Die Haupteffekte jener Informationstechnologie, die noch nicht auf Maschinelernen beruht, haben sich bereits in einer engen Nische bei Routine-tätigkeiten, hoch strukturierten und sich wiederholenden Aufgaben eingestellt«, schreiben sie. In den kommenden Jahren werde nun ein viel breiteres Spektrum von Aufgaben automatisiert oder durch Maschinen angereichert

tain types of cancer or pneumonia as well or even better than medical specialists, their ability to explain how they established their diagnoses is extremely low. However, Brynjolfsson and Mitchell concede that it is sometimes also very difficult for humans to explain how they solve certain tasks, such as that of identifying words from the sounds they hear.

6. **There is a tolerance for error and no need for provably correct or optimal solutions.** Almost all machine learning systems derive their solutions statistically. It is therefore rarely possible to train them to one-hundred-percent accuracy. Even the best speech or object recognition systems or clinical diagnosis computers do make mistakes. A certain tolerance of the system to errors is therefore a fundamental criterion when it comes to the application of learning algorithms.
7. **The phenomenon or function to be learned should not change rapidly over time.** In general, machine learning algorithms work well if they are applied for tasks resembling the training examples. Changes in application will consequently require retraining. Success therefore depends on the rate of change—relative to the rate of acquisition of new training data. The two scientists state e-mail spam filters as a positive example: the number of unsolicited e-mails is relatively high compared to how fast the type of spam mails changes.
8. **No specialized dexterities, physical skills, or mobility is required.** Robots are still rather clumsy compared to humans when facing unstructured environments and tasks. According to Brynjolfsson and Mitchell, this is not so much a deficit of machine learning but rather a consequence of the state of the art of robot technology.

Only certain parts of a job are affected

Based on their analysis, Brynjolfsson and Mitchell have drawn several general conclusions as to the role of artificial intelligence in the future world of labor. “The main effects of pre-machine-learning information technology have been on a relatively narrow swath of routine, highly structured, and repetitive tasks,” they maintain, convinced that in the years to come a much broader spectrum of tasks will be automated or augmented by machines. “This includes tasks for which humans are unable to articulate a strategy, but where statistics in data reveal regularities that entail a strategy.”

werden. »Das beinhaltet Aufgaben, bei denen der Mensch nicht in der Lage ist, eine Strategie zu artikulieren, bei denen aber statistische Zusammenhänge in Daten Regelmäßigkeiten offenbaren, die eine Strategie beinhalten.«

Daraus folgt, dass sich die Einsatzbereiche von Maschinen ändern werden und dass es daher irreführend wäre, vergangene Trends einfach in die Zukunft zu extrapolieren. Jobs bestehen in der Regel aus einer Reihe von verschiedenen, aber zusammenhängenden Aufgaben. »In den meisten Fällen sind wahrscheinlich nur einige dieser Aufgaben für Maschinenlernen geeignet – und das sind nicht unbedingt jene, die mit früheren Technologien einfach zu automatisieren waren.« Die Forscher nennen einige Beispiele: Einer lernenden Maschine könne beispielsweise beigebracht werden, Anwälten zu helfen, Dokumente zu klassifizieren, die für einen Fall potenziell relevant sind. Eine Maschine könne jedoch schwer Zeugen interviewen oder eine Strategie entwickeln, um in einer Gerichtsverhandlung Recht zu bekommen. Oder: Künstliche Intelligenz könne Krebsdiagnosen stellen, kann aber nicht wie ein Arzt erklären, warum sie zu diesem Ergebnis gekommen ist. Auch wenn es um fachlichen Austausch mit anderen Ärzten oder um eine emotional herausfordernde Kommunikation mit Patienten gehe, werde man weiterhin medizinisches Fachpersonal benötigen.

Wo hier die Grenze verläuft – ob Maschinen also auch soziale Kompetenzen bekommen oder ob sie so etwas wie Kreativität entwickeln können –, darüber wollen die beiden Wissenschaftler keine allgemeinen Aussagen treffen. Sicher sei jedenfalls, dass die verbleibenden Aufgaben, die von Maschinen nicht übernommen werden können, wichtiger und wertvoller werden: Die menschlichen Fähigkeiten werden also aufgewertet. So wird etwa die Rolle des Menschen wichtiger, klare Ziele zu definieren und die richtigen Fragen zu stellen (die Maschinen dann oft besser beantworten können).

Veränderungen brauchen ihre Zeit

Brynjolfsson und Mitchell betonen, dass die Auswirkungen von künstlicher Intelligenz auf die Arbeitswelt nicht nur von technologischen Faktoren, sondern sehr stark auch von rein ökonomischen Zusammenhängen bestimmt werden. Veränderungen werden jedenfalls nicht von heute auf morgen geschehen. Die Einführung und Verbreitung von Technologien dauert oft Jahre oder Jahrzehnte – weil parallel dazu in vielen Bereichen Änderungen notwendig sind: etwa Änderungen von Produktionsprozessen, des Organisationsdesigns, von Geschäftsmodellen, Lieferketten, rechtlichen Rahmenbedingungen bis hin zu kulturellen Erwartungen. Im Detail seien diese Prozesse kaum zu überblicken: »Solche Komplementaritäten sind in modernen Organisationen und Ökonomien so allgegenwärtig, wie sie subtil und schwer zu identifizieren sind«, so die beiden Forscher. ✕

Quelle: Erik Brynjolfsson und Tom Mitchell, »What can machine learning do? Workforce implications«, in: *Science*, Bd. 358, Nr. 6370 (22.12.2017), S. 1530–1534; DOI: 10.1126/science.aap8062.

This suggests that the scope of machine applications will change and that it would therefore be misleading to extrapolate past trends into the future. Jobs usually consist of a number of individual, albeit interconnected tasks. “In most cases, only some of these tasks are likely to be suitable for machine learning, and they are not necessarily the ones that were easy to automate with previous technologies.” The two scientists give a number of examples—such as that of a learned system that could help lawyers identify potentially relevant documents for a case, although it would hardly be possible for a machine to interview witnesses or develop a strategy fit to prevail in a court hearing; or that of artificial intelligence capable of diagnosing cancer, although the system, unlike a physician, cannot possibly explain how it arrived at a particular result. Medical experts will also continue to be indispensable as far as the exchange among experts or the emotionally challenging task of communicating with patients is concerned.

In the end, however, the two researchers refrain from drawing a line here—whether machines can actually acquire social competencies or develop something like creativity—and from making any general predictions in this respect. In any case there can be no doubt that the remaining tasks machines will be unable to perform will become more relevant and more valuable: human abilities will thus be upgraded. The role of humans will become more important when it comes to defining goals clearly and asking the right questions (which machines will then often be able to solve more efficiently).

Change needs time

Brynjolfsson and Mitchell emphasize that the implications of artificial intelligence for the workforce will not only be affected by technological factors but, to a large extent, also by purely economic considerations. Change will certainly not happen overnight. The introduction and diffusion of technologies often take years if not decades—because they need to go hand in hand with complementary changes in many other areas: such as in production processes, organizational design, business models, supply chains, legal frameworks, and even cultural expectations. These processes can barely be foreseen in detail: “Such complementarities are as ubiquitous in modern organizations and economies as they are subtle and difficult to identify,” the two scientists say. ✕

Source: Erik Brynjolfsson and Tom Mitchell: “What can machine learning do? Workforce implications,” *Science*, vol. 358, issue 6370 (December 22, 2017), 1530–1534; DOI: 10.1126/science.aap8062.



G
K

D
l
w
S
Z
S
e
d
E
S
w
a
d
a
B
a
5
t
z
G
K
g
T
h
a
a
t
i
r
b
g
e
t
i
c
m
t
c
i
c
i
s
t
T
c
g

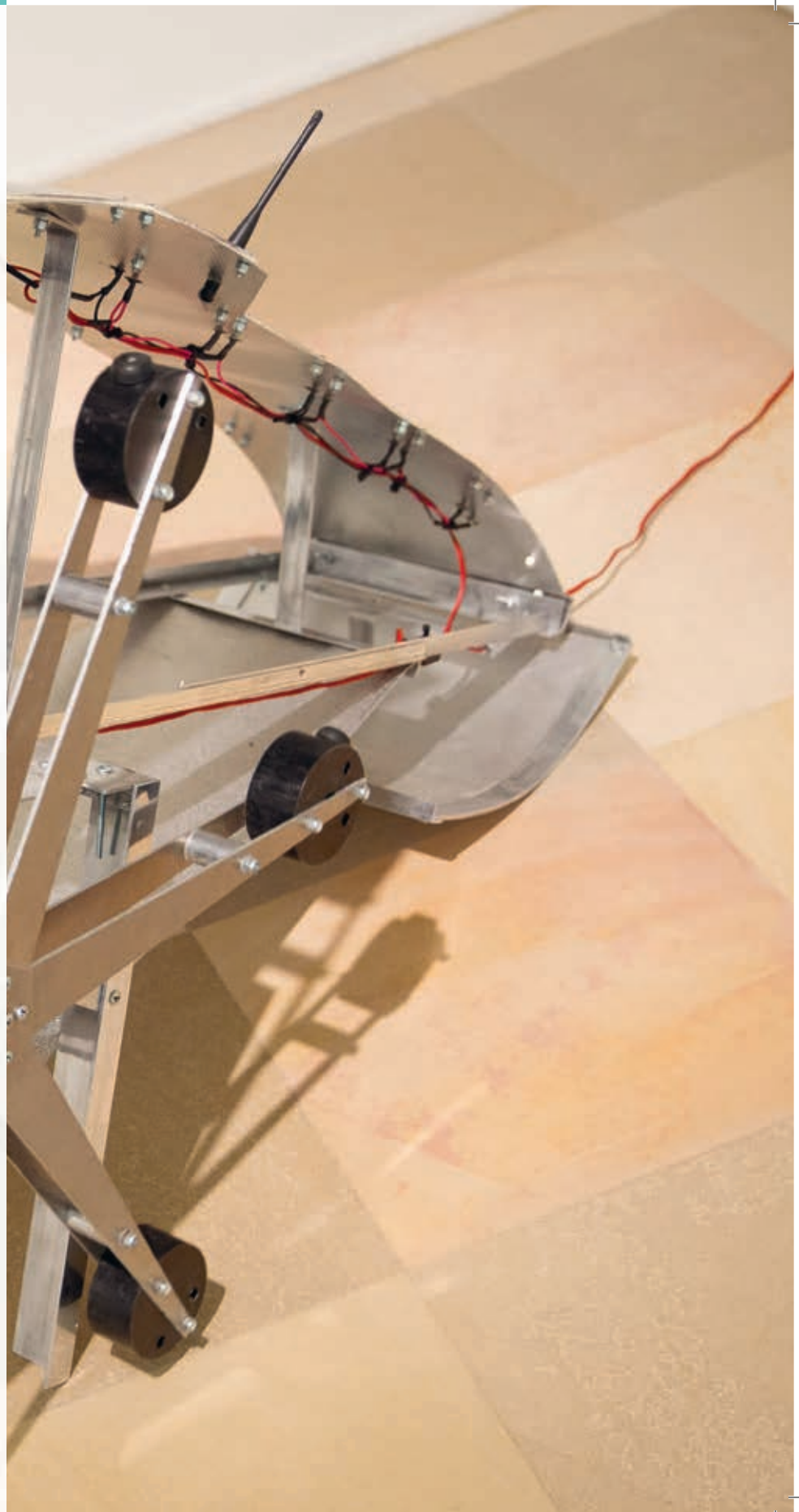
—
©
tr

Glaciator Joaquín Fargas (AR)

Der »Glaciator«-Roboter lässt Gletscher wieder wachsen, indem er den Schnee zu Firn (die Zwischenstufe zwischen Schnee und Gletschereis) verdichtet und damit den Prozess der Eisbildung vorantreibt. So können die Gletscher wieder jene Eismassen aufbauen, die aufgrund des Klimawandels abgeschmolzen sind. Besonders in der Antarktis hat in den letzten 50 Jahren eine dramatische Gletscherschmelze stattgefunden. Grund dafür sind der Klimawandel und die globale Erwärmung.

The "Glaciator" robot helps glaciers to grow again by compacting and crystallizing the snow into firn (the intermediate state between snow and glacier ice), thus accelerating the ice formation process. Glaciers can thus grow the ice mass again, which they have lost due to climate change. Especially in Antarctica the ice mass has diminished dramatically in the last fifty years. The reasons for this are climate change and global warming.

© Ars Electronica/
tom mesic



Der digitale Stall: Wie Roboter in unser Leben drängen

Maschinen, die gewisse Arbeiten für uns übernehmen, kommen in immer mehr Bereichen zum Einsatz. Die Schlüssel dafür sind zum einen Fortschritte in der Sensorik und Elektronik und zum anderen vor allem die Intelligenz, die im Hintergrund viele Fähigkeiten von Robotern erst ermöglicht. Ein kurzer Rundblick von der Landwirtschaft bis auf die Straße.

Neulich im Waldviertel: Beim »Kuhfrühling« in der Versteigerungshalle Zwettl wurden zwei Tage lang Kühe aller Altersklassen fach- und sachgerecht präsentiert, Wertungen durch Preisrichter inklusive. Dem Publikum, das die Arena bis auf den letzten Platz und mit Volksfeststimmung füllte, gefiel es offenkundig. Es wurde auch einiges geboten: Man konnte nicht nur die Tiere, allesamt vor Gesundheit strotzende Exemplare, aus der Nähe bestaunen, es wurden auch die allerneuesten Innovationen aus der Milchviehwirtschaft gezeigt. Vollautomatische Melkstände, die jedes Tier individuell nach dessen Bedürfnissen behandeln und »abarbeiten«, hat man ja schon gesehen – aber ein Fütterungsroboter war wirklich etwas Neues. Dabei handelt es sich um ein tonnenförmiges Gerät auf Rädern, das sich selbsttätig beim Silo nachfüllt, das Futter in seinem Inneren mischt, zu den verschiedenen Gruppen von Kühen im Stall (Milchkühe, Jungvieh, trockenestellte Kühe) fährt und dort die jeweils optimale Futterration ablädt. Um das erledigen zu können, hat der Roboter zahlreiche Sensoren eingebaut, die unter anderem prüfen, ob der Fahrweg frei ist, den Abstand zum Fressgitter registrieren und die abgegebene Futtermenge dosieren und dokumentieren – und ermitteln, ob irgendwo noch altes, nicht gefressenes Futter liegt. Der Bauer kann per App auf dem Smartphone individuelle Fütterungspläne für die verschiedenen Tiere zusammenstellen und so mit minimiertem Arbeitsaufwand eine höhere Milchleistung der Kühe erzielen.

Wer nun glaubt, dass ein derartiger Futterroboter nur etwas für einige wenige »nerdige« Jungbauern sei, der irrt gewaltig: Die Landwirtschaft ist nämlich alles andere als in Traditionen verhaftet, sondern vielmehr eine äußerst innovative Branche. Auf größeren Feldern, etwa im Marchfeld östlich von Wien, ist es längst gang und gäbe, dass Traktoren GPS-gesteuert ihre Furchen ziehen und Erntemaschinen die Feldfrüchte nach einem optimierten Plan aussäen und einbringen. Moderne Geräte ermitteln während der Fahrt die Bodenqualität oder den Zustand der aufgehenden Pflanzen und dosieren gemäß diesen Messwerten die nötigen Dünger- oder Pestizidgaben.

Automatische Salaternte

Der Mensch als Fahrer der Maschinen ist dabei noch notwendig. Aber in Labors und Versuchsbetrieben sind auch schon autonome Agrarmaschinen startklar, die ohne Zutun des Menschen agieren. Ein Beispiel dafür ist der Feld-

Digital Barns: How Robots Are Pushing into Our Lives

Machines taking over certain tasks for us have come to be used in more and more fields. The key prerequisites for this are advances in sensor technology and electronics as well as, above all, the background intelligence enabling numerous robot capabilities. The following presents a short tour d’horizon from agriculture to the streets.

Recently in the Waldviertel region of Lower Austria: cattle of all age categories, complete with the judges’ scores, were presented professionally and appropriately at the “Cattle Spring” event in the auction hall of Zwettl. Filling the arena to bursting and creating a carnival atmosphere, the public obviously loved it. No wonder, considering the range of offers: visitors could not only admire the animals, all of them the picture of health, at close range but also make themselves familiar with the most recent innovations in dairy farming. They might have already seen fully automated milking parlors dealing individually with each animal according to its demands—yet a cattle feeding robot was really something new to them. The drum-like device on wheels automatically refills itself from the silo, mixes the feed, and drives to the different groups of cows in the barn (dairy cows, young stock, dried-off dairy cows) where it unloads the respective optimum feed ration. The robot is furnished with numerous sensors to be able to fulfil these tasks. These sensors check whether the way is free, register the distance to the feed grid, dose and document the feed provided—and find out whether there is any old feed left that has not been eaten. Using an app on his smartphone, the farmer can design feeding plans for different animals and thus increase the milk production level with a minimized amount of work.

Who considers such a feeding robot a device for just a few nerdy young farmers is seriously mistaken: agriculture is anything but caught up in traditions; it is an extremely innovative sector. On larger fields such as those of the Marchfeld area east of Vienna, it has been common practice for a long time to rely on GPS-controlled tractors to plough furrows and on harvesters for seeding and gathering in the crop. While on the move, modern machines determine the soil quality and the status of sprouting plants and define the required fertilizer and pesticide doses according to these data.

Automated lettuce harvesting

Machines like these still depend on human drivers. Yet autonomous agricultural machines operating without human help are ready to start in certain laboratories and trial run farms. The field robot FRANC, developed by Austrian scientists under the overall control of Vienna’s University of Technology, provides a striking example. This robot navigates automati-

roboter FRANC, der von österreichischen Forschern unter Federführung der TU Wien entwickelt wurde und wird. Dieser kann automatisch über das Feld navigieren, erkennt Hindernisse, findet selbsttätig die angebauten Pflanzenreihen und kann beispielsweise mit Hackrädern zur Unkrautbekämpfung ausgestattet werden. Für solche Maschinen sind viele Einsatzgebiete denkbar – etwa dass sie sich völlig selbsttätig um Salat kümmern: Dabei wird bei jeder Pflanze einzeln betrachtet, wie es ihr geht, ob sie Wasser oder Dünger benötigt oder ob sich Schädlinge breitmachen. Wenn zudem das Vermarktungssystem meldet, dass gerade große Nachfrage herrscht und ein guter Preis erzielt werden kann, werden automatisch die reifsten Salatköpfe geerntet und zum Sammelplatz gebracht.

All diesen Systemen sind zwei Dinge gemein: Erstens nehmen sie dem Menschen Arbeit ab, und zweitens können sie diese Arbeit in einer zuvor nicht bekannten Präzision erledigen, sodass beispielsweise nur so viel Dünger ausgebracht wird wie unbedingt nötig. Das spart Ressourcen und schont überdies die Umwelt. Diese beiden Eigenschaften gelten für alle Roboter, wobei das Beispiel der Landwirtschaft zeigt, dass Roboter in vielen Fällen nichts mit dem vorherrschenden Bild in unserem Kopf zu tun haben, demzufolge immer große Geräte mit Greifarmen am Werk sind. Dieses Bild repräsentiert Industrieroboter, etwa in einer Auto- oder in einer Elektronikfabrik. Doch die moderne Robotik geht weit darüber hinaus. Heute werden auch automatische Staubsauger oder Rasenmäher unter den Begriff Roboter subsumiert – und mittlerweile auch nichtmaterielle Systeme wie etwa Sprachautomaten, die als »Bots« (Kurzform von »Roboter«) bezeichnet werden. Als Sonderform müssen androide Roboter gelten, also autonome Systeme, denen Menschengestalt verliehen wurde. Wie etwa »Sophia«, die beim heurigen Forum Alpbach ihren großen Auftritt hat. Der in Hongkong gefertigte Roboter der allerneuesten Generation spricht zum Beispiel (verständliches) Englisch und kann einfache Phrasen auf Deutsch wiederholen.

Kollaborative Roboter

Herkömmliche Roboter waren dafür konzipiert, in einem definierten und abgeschlossenen Bereich (typischerweise hinter Schutzgittern) festgelegte und sich ständig wiederholende Tätigkeiten zu verrichten, etwa einen Gegenstand zu nehmen, zu bewegen oder ihn zu verschweißen. Dabei kam es zu keiner direkten Interaktion mit dem Menschen. Diese Konzeption ändert sich nun: Roboter werden zunehmend dafür ausgelegt, mit Menschen zusammenzuarbeiten und sie im Sinn einer Arbeitsteilung bei ihren Aufgaben zu unterstützen – man nennt diese kollaborativen Roboter »Cobots«. Dazu benötigen die Geräte zwei Dinge: erstens Sinnesorgane, mit denen sie ihre Umgebung wahrnehmen, und zweitens eine gewisse »Intelligenz«, mit der sie die Sensordaten verarbeiten und daraus sinnvolle Schlüsse ziehen, die sich am Ende in ihrer Tätigkeit widerspiegeln.

Entscheidend ist in vielen Fällen, dass der künstlichen Intelligenz ein ausreichend großer und valider Datenbestand zur Verfügung steht, anhand dessen die Systeme lernen können – etwa sich wiederholende Muster zu erkennen.

cally across the fields, recognizes obstacles, finds the rows of plants to be cultivated, and can be furnished with wheel weeders. Many areas of application are envisaged for such machines. They might, for instance, fully automatically take care of lettuce: examining each plant individually, ascertaining its condition, determining whether it needs water or fertilizer, or whether it is attacked by pests. When the marketing system reports a great demand and a good price to be achieved, the machine automatically harvests the ripest heads of lettuce and transports them to the gathering place.

All these systems share two qualities: they save people a lot of work and accomplish their tasks in a hitherto unheard-of precision, so that only the absolutely necessary amount of fertilizer will be spread, for instance. This approach taps fewer resources and saves the environment. The two features described hold true for all robots. The example of agriculture demonstrates that many robots have nothing to do with the large machines equipped with gripper arms dominating our imagination. This picture reflects industrial robots such as those in an automobile or electronics factory. Modern robotics goes far beyond this, however. Today, automatic vacuum cleaners and lawnmowers tend to be subsumed under the term “robot,” too. In the meantime, this also applies to non-material systems such as language machines, which are called “bots” (short for “robots”). Android robots, i. e., autonomous systems designed to look like humans, constitute a special form. “Sophia,” who will make her big appearance at this year’s Forum Alpbach, is such an android. Made in Hong Kong, this robot of the very latest generation speaks (understandable) English and can repeat simple phrases in German.

Collaborative robots

Traditional robots were designed to perform defined and continuously repeated activities in a demarcated and closed area (typically behind safety barriers), e. g., pick up an object, move it, or weld something. These activities did not include any direct interaction with humans. This concept is changing now: robots are increasingly built to collaborate with humans and to support them in performing their tasks along the lines of a division of labor. These collaborative robots are called “co-robots” or “cobots.” They require two things: sensors to discern the world around them and a certain “intelligence” that enables them to process the sensory data and draw conclusions from them that will finally be reflected in their activities.

In many cases, artificial intelligence decisively depends on the availability of a sufficiently large and valid set of data that provides systems with the possibility to learn—to recognize repetitive patterns, for instance.

Robots on the street

It will be necessary to improve the computing power, sensor technology, and data analysis capacity especially for robots not operating in well-structured environments but in the real world: at hotel check-ins, in

Roboter auf der Straße

Eine Verbesserung der Technologien hinsichtlich Rechenleistung, Sensorik und Datenanalyse wird vor allem auch für jene Roboter benötigt, die nicht mehr in wohlstrukturierten Umgebungen tätig sind, sondern »draußen« in der realen Welt – sei es als Helfer beim Einchecken in einem Hotel, sei es für Pflegedienstleistungen, beim Rasenmähen oder beim Futterverteilen im Stall. Das beste Beispiel sind autonome Autos (für die sich immer mehr der Begriff »Roboter-Autos« durchsetzt). Diese müssen anhand von (durch Ultraschall-, Laser-, Infrarot- und Kameravorrichtungen usw. gelieferte) Sensordaten ein Bild der Verkehrslage gewinnen, dieses mit zahlreichen anderen Informationen wie etwa Wettervorhersagen, Stauwarnungen, Meldungen von anderen Fahrzeugen auf der Straße oder Angaben zum Ladezustand der Batterie verknüpfen und daraus in Echtzeit die richtigen Handlungen ableiten. Heutige Fahrerassistenzsysteme wie Spurhalteassistenten, Einparkhilfen oder Notbremssysteme können das in ihren Bereichen schon ziemlich gut. Vollautonomes Fahren erfordert aber noch eine deutliche Weiterentwicklung der Technologien.

Eines der Hauptprobleme dabei ist, dass all das noch von unzähligen äußeren und vielfach unvorhersehbaren Faktoren beeinflusst wird. Wie komplex die Welt ist, in der sich autonome Roboter bewegen, mussten kürzlich auch Ingenieure von Uber zur Kenntnis nehmen: In Tempe in Arizona kam es Mitte März 2018 zum ersten Todesopfer durch ein autonomes Auto. Eine Frau, die ein Rad schiebend die Straße querte, wurde von dem Auto erfasst. Die örtliche Polizei meinte zwar, dass auch ein Mensch am Steuer diesen Unfall nicht hätte verhindern können. Dennoch kommt nun die große Vision hinter dem autonomen Fahren, nämlich nicht nur den Komfort, sondern auch die Verkehrssicherheit drastisch zu erhöhen, gehörig unter Druck. .

Das Beispiel zeigt jedenfalls, dass die Technologie ihren Zenit noch lange nicht erreicht hat und dass es bei der fortschreitenden Durchdringung unserer Lebenswelt mit Robotern noch viele Überraschungen geben wird – so manche unliebsame, aber wohl auch viele positive. Bei der Hannover Messe, einer der wichtigsten Industriemessen der Welt, wurden im April des Jahres beispielsweise Roboter gezeigt, die Weißbier ausschenken oder Tischfußball spielten – und bei jeder Partie dazulernten. Eine große Zukunft wird Medizin-Robotern vorhergesagt, etwa bei minimalinvasiver Chirurgie, bei der Positionierung von Röntgensystemen oder bei der robotergeführten Strahlentherapie.

Die Zahl der Roboter steigt jedenfalls rasant: Der International Federation of Robotics zufolge wurden im Vorjahr weltweit rund 346.000 Industrieroboter verkauft, um 18 Prozent mehr als im Jahr zuvor. Bis 2020 soll der Absatz laut Prognosen der IFR weiter auf 521.000 steigen. ✖

Dieser Beitrag ist eine erweiterte Fassung des Artikels »Wie Roboter in unser Leben drängen«, der im *UNIVERSUM Magazin*, Ausgabe Mai 2018, erschienen ist.

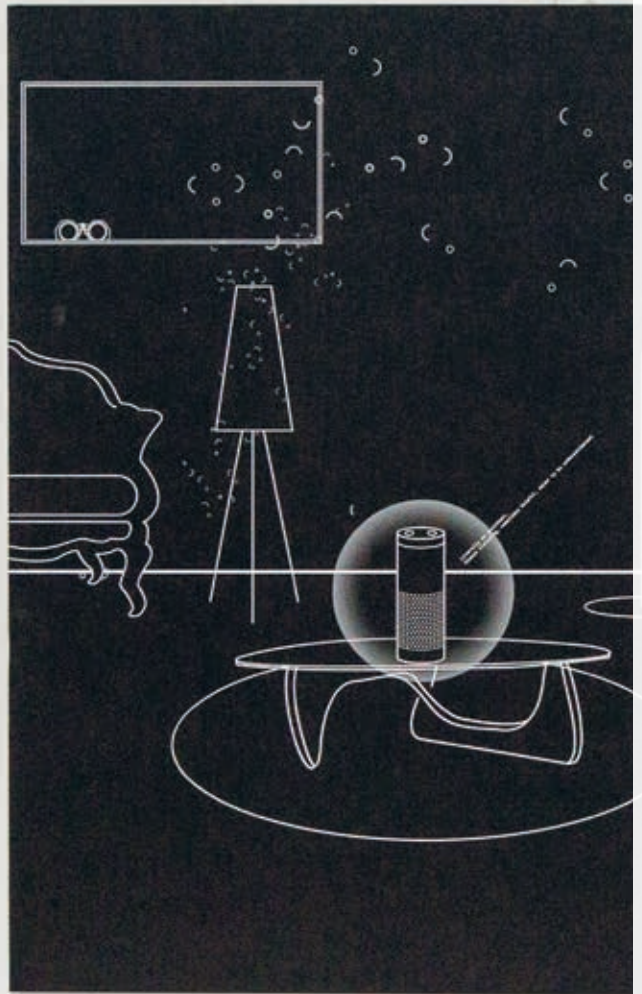
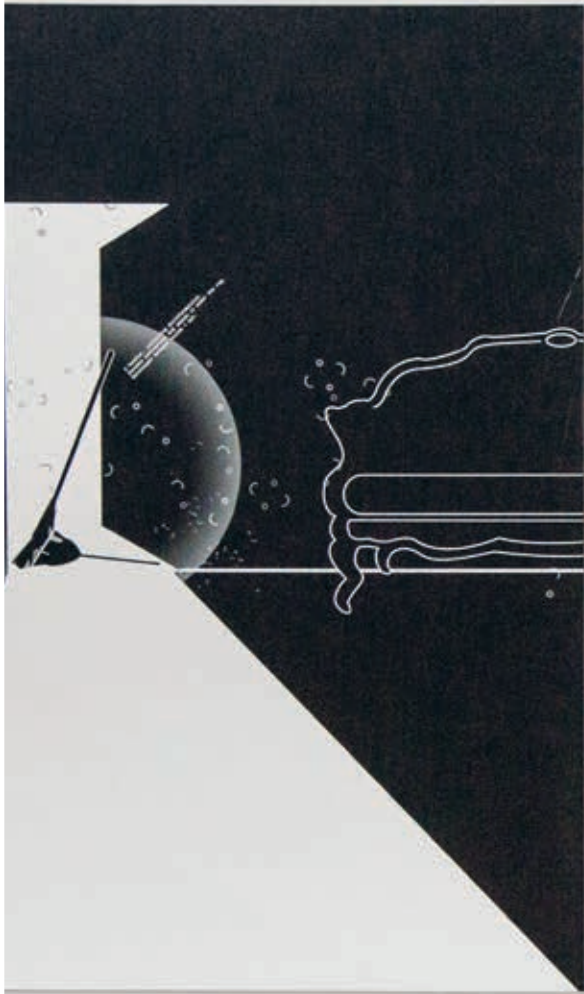
nursing services, for cutting the lawn, or for distributing fodder to cattle. The best example are autonomous cars (increasingly called “robot cars”). Relying on sensory data (provided by ultrasound, laser, infrared, and camera devices, etc.), such cars generate a comprehensive picture of the traffic situation. They interlink these data with numerous other bits of information such as weather forecasts, traffic congestion warnings, reports on other vehicles on the street, or the battery’s state of charge and derive the right activities in real time. Today’s advanced driver assistance systems such as lane departure warning systems, parking sensors, or automatic emergency breaking systems are already quite good at this. Fully autonomous driving, however, definitely requires further developments of these technologies.

One of the crucial problems in this context is the influence of innumerable external and often unforeseeable factors. Uber engineers were recently confronted with the complexity of the world in which autonomous cars move: an accident involving an autonomous car claimed the first life in Tempe, Arizona, in mid-march 2018. The car struck a woman walking her bike across the street. Local police authorities maintained that it would have been equally impossible to avoid the collision for a human behind the wheel. Nevertheless, the great vision behind autonomous driving, i.e., a drastic increase of both comfort and road safety, has been exposed to quite a lot of pressure by this incident.

Be that as it may, the example shows that the relevant technology has clearly not reached its zenith and that the progressive pervasion of our environment with robots will hold numerous surprises—some unwelcome ones, but also many positive ones. The Hanover Fair, one of the most important industrial fairs in the world, amazed its visitors with robots that served wheat beer or played foosball—and learned something with each match. Medicine robots are expected to have a great future, such as in minimally invasive surgery, in the positioning of X-ray systems, and in robot-assisted radiotherapy.

The number of robots is rapidly increasing: according to the International Federation of Robotics, about 346,000 industrial robots were sold worldwide in 2017, which amounts to 18 percent more than the year before. IFR predicts that the number of industrial robots being supplied by 2020 will be around 521,000. ✕

This text is an extended version of the article “Wie Roboter in unser Leben drängen,” which was published in *UNIVERSUM Magazin*, May 2018.



10F6



Evidence



○



20F6

Evidence



Several participants in the workshop were asked to describe their own work. A participant from the University of California, Berkeley, described their work as follows: "I am a data scientist and I work on the intersection of machine learning and human-computer interaction." "I am a data scientist and I work on the intersection of machine learning and human-computer interaction."

"NB: NOTE: how to generate? in case of crime, who is at fault: code? human? device?"

**Broom:
Coding +
Decoding**

The system will continue to be updated and improved. It will be used to monitor the security of the system and to detect any suspicious activity.



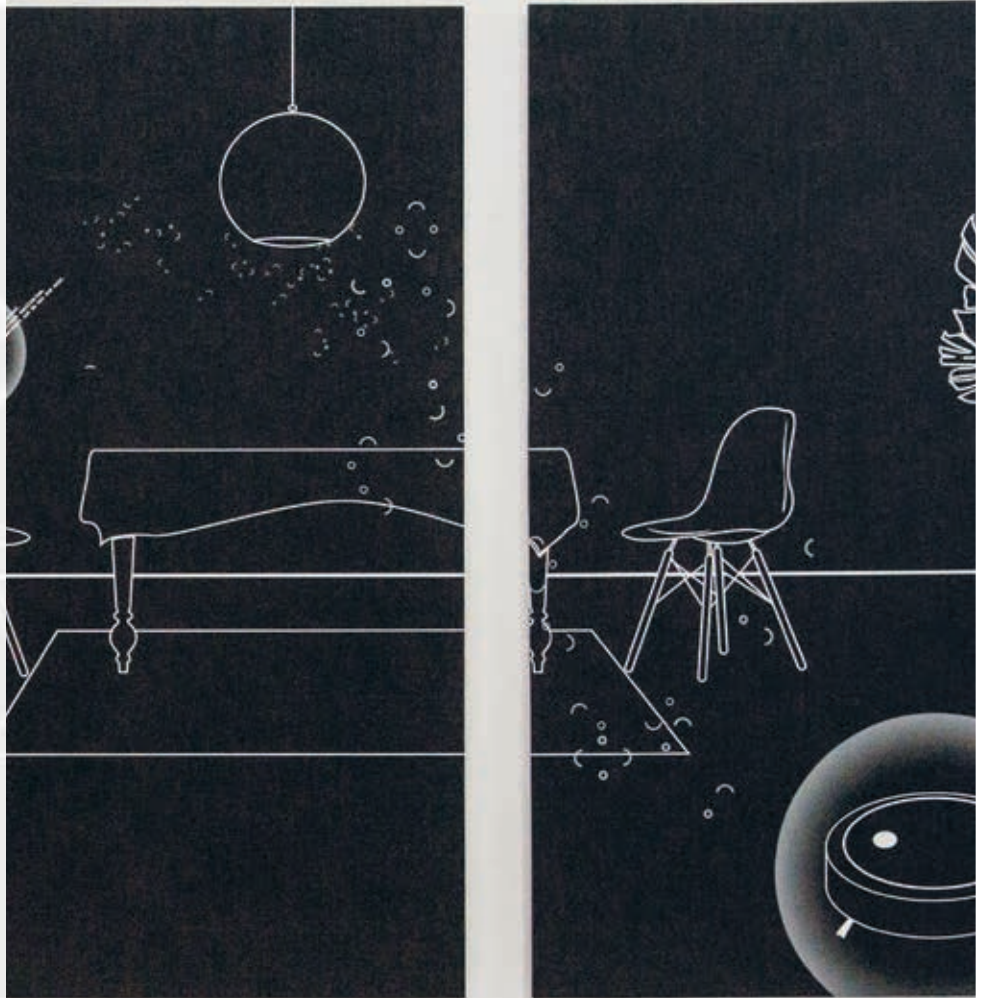
Evidence

The Internet of Enlightened Things
ArtCenter College of Design,
Pasadena (us)

»The Internet of Enlightened Things« ist eine Sammlung von Projekten, die untersuchen, welche Implikationen und Chancen sich ergeben, wenn wir unser Leben – willentlich oder nicht – mit immer »intelligenteren« Objekten und Systemen teilen.

“The Internet of Enlightened Things” is a collection of projects that explore the implications and opportunities of sharing our lives—willingly or not—with ever more “intelligent” objects and systems.

© Ars Electronica/
 Florian Voggeneder



40F6

Evidence

The process of confirming the existence, data flow, and control of the Internet of Things



How do we know the appliances are connected to the Internet? How do we know the appliances are connected to the Internet? How do we know the appliances are connected to the Internet?

How do we know the appliances are connected to the Internet? How do we know the appliances are connected to the Internet? How do we know the appliances are connected to the Internet?

Evidence

**Roomba:
 Spatial Mapping
 + Navigation**

Georg Langs

Präzisionsmedizin und künstliche Intelligenz



Künstliche Intelligenz und vor allem maschinelles Lernen versprechen den Zugang zur Präzisionsmedizin zu verändern. Das Verarbeiten und Erkennen komplexer Muster, um Spuren von Krankheiten zu entdecken, Diagnosen zu bewerten oder Ergebnisse vorherzusagen, sind vorrangige Einsatzformen künstlicher Intelligenz zur Lösung bekannter Probleme. Gleichzeitig hat die Verfügbarkeit zusehends effektiver Daten zur Entstehung neuer Ansätze geführt, die Diagnose- und Therapieentscheidungen unterstützen. Wir stehen allerdings noch immer am Anfang der Reise. Erfolg bedarf der interdisziplinären Zusammenarbeit von Medizin und Informatik.

Von der Unterstützung alltäglicher Tätigkeiten bis zu den Bereichen Automatisierung, Industrie und Mobilität gewinnt künstliche Intelligenz (KI) rasch an Bedeutung. Sie steht im Begriff, zu einer Fähigkeit zu werden, die Aufgaben erfüllt, für die menschliche Intelligenz als notwendig erachtet wird. Ein besonders vielversprechender Ansatz, durch den Rückgriff auf eine Strategie des Menschen zu KI zu gelangen, ist maschinelles Lernen: Statt sozusagen händisch große Regelsysteme zu entwerfen, werden Computer in die Lage versetzt, aus Beispielen zu lernen. Dieser Ansatz eignet sich besonders für Anwendungen in der Medizin, da man in diesem Bereich Regeln häufig a priori festzulegen versucht. Auf der Grundlage einer großen Anzahl von Beobachtungen kann maschinelles Lernen komplexe Zusammenhänge und Merkmale identifizieren, die für die Unterstützung von Entscheidungen und das Verständnis der diesen zugrundeliegenden Mechanismen erforderlich sind.

Wird maschinelles Lernen für Diagnose und Prognose eingesetzt, sieht man sich allerdings mit Herausforderungen konfrontiert, die ganz anders geartet sind als die in anderen Bereichen, in denen sich maschinelles Lernen als sehr erfolgreich erwiesen hat, wie etwa bei der Gesichtserkennung oder dem Verstehen von Sprache. Eines der Hauptprobleme ist, dass die Signaturen, die Experten ihres Fachs zu erkennen und zu benutzen gelernt haben, um in der Behandlung einzelner Patienten weitreichende Entscheidungen zu steuern, im Vergleich zur grundsätzlichen natürlichen Verschiedenheit von Menschen äußerst subtil sind. Zu Daten eines Maßstabs zu gelangen, der in der natürlichen Bildverarbeitung die Norm ist, ist in der Medizin bis heute schwierig. Während man mit datengesteuerten Ansätzen vielversprechende Resultate erzielt, besteht schließlich ein wichtiger Aspekt der Anwendung von KI in der Medizin darin, dass es die zugrundeliegenden Mechanismen zu verstehen gilt, damit die Ergebnisse maschinellen Lernens zum Ausgangspunkt für bestimmte auf Hypothesen beruhende Forschungen werden können.

Georg Langs leitet das Computational Imaging Research Lab (www.cir.meduniwien.ac.at) und ist a.o. Professor am Department of Biomedical Imaging and Image-guided Therapy der Medizinischen Universität Wien sowie Chief Scientist der contextflow GmbH (www.contextflow.com), einem Ableger der Medizinischen Universität, der von Deep-Learning-Prozessen gesteuerte Suchalgorithmen für medizinische Bilddaten entwickelt. Langs hat an der Technischen Universität Wien Mathematik und an der Technischen Universität Graz Informatik studiert und als Postdoktorand am Laboratory für angewandte Mathematik der École Centrale Paris und Forschungsassistent am Computer Science and Artificial Intelligence Laboratory des Massachusetts Institute of Technology (MIT) gearbeitet, dem er noch immer verbunden ist.

Georg Langs is Head of the Computational Imaging Research Lab (www.cir.meduniwien.ac.at), Associate Professor at the Department of Biomedical Imaging and Image-guided Therapy at the Medical University of Vienna, and Chief Scientist at contextflow GmbH (www.contextflow.com), a MedUni spin-off that is developing deep-learning-driven search algorithms for medical imaging data. Langs studied mathematics at the Vienna University of Technology and computer science at the Graz University of Technology. He worked as a postdoc at the Applied Mathematics Laboratory of the École Centrale Paris and as a research scientist at the Computer Science and Artificial Intelligence Laboratory of the Massachusetts Institute of Technology (MIT), with which he is still affiliated.

Georg Langs

Precision Medicine and Artificial Intelligence

Artificial intelligence and in particular machine learning are promising to change our approach towards precision medicine. The processing and recognition of complex patterns to detect traces of disease, to estimate diagnosis, or to predict outcome are prime uses of artificial intelligence to solve problems we know. At the same time, the availability of increasingly powerful algorithms has led to new approaches supporting diagnosis and treatment decisions. Yet we are still at the beginning of this journey. Interdisciplinary work among medicine and computer science is a necessity for success.

Artificial intelligence (AI) is rapidly gaining relevance in fields ranging from supporting our day-to-day activities to automation, industry, and mobility. It is moving toward the ability to perform tasks that are viewed to require human intelligence. Machine learning is a particularly promising approach to achieve AI by exploiting a human strategy. Instead of designing large systems of rules by hand, as it were, we enable computers to learn from examples. This strategy is particularly well suited for applications in medicine, since often we struggle to define rules a priori in this field. Based on large amounts of observations, machine learning can identify complex relationships and characteristics necessary for decision support and the advance of our understanding of underlying mechanisms.

However, when applying machine learning to diagnosis or prognosis in medicine, we face challenges unlike those associated with other fields where machine learning has proven immensely successful, such as identifying locations in photographs, recognizing faces, or understanding language. A primary challenge is that the signatures medical experts have learned to detect and exploit to guide far-reaching decisions in the treatment of individual patients are extremely subtle compared to the substantial natural variability among humans. To date, obtaining data of a scale that is standard in natural image processing is difficult in medicine. Finally, while data-driven approaches achieve promising results, an important aspect of AI in medicine is that there is a need to understand underlying mechanisms, and so the results of machine learning become a starting point for specific hypothesis-driven research.

Ein Gebiet, das von KI profitieren wird, ist die Präzisionsmedizin, die darauf abzielt, Patienten individuell zu versorgen. Um das möglich zu machen, werden individuell unterschiedliche Krankheitsverläufe und Reaktionen auf die Behandlung in Betracht gezogen. Die Präzisionsmedizin betrachtet eine Krankheit und die entsprechende Therapie nicht als singuläre Kategorie, sondern als komplexes Muster individueller Merkmale des Patienten, des Verlaufs seiner Krankheit und der verschiedenen möglichen Reaktionen auf die Bandbreite verfügbarer Behandlungen.

Es geht vor allem darum, die Muster und Beziehungen zwischen diesen komplexen Beobachtungen und der Entscheidung für eine bestimmte Behandlung zu erkennen und daraus Nutzen zu ziehen. Obgleich KI nur anfangs zu vielversprechenden Ergebnissen geführt hat, hat Deep Learning zu einiger positiver Aufregung Anlass gegeben, da sich mit diesem Ansatz bemerkenswert genaue Vorhersagen treffen lassen. Gelitten hat der Ruf von Deep Learning darunter, dass man diese Technik maschinellen Lernens als Black Box dargestellt hat, die zwar gut funktioniert, aber schwer zu deuten ist. In jüngster Zeit hat man Methoden vorgestellt, tiefe neuronale Netze auf Merkmale zu untersuchen, die Vorhersagen zulassen und zu verstehen erlauben, welche Teile der Daten nützliche Informationen enthalten, wenn es darum geht, Modelle der Beziehung zwischen Beobachtung und Vorhersageziel zu entwerfen.

Klarer wird, dass KI in der Präzisionsmedizin verschiedene Rollen übernehmen wird: Die Palette reicht von der Automatisierung sich wiederholender Aufgaben, welche die Verarbeitung von Informationen wie etwa Bildern erfordern, bis zur Erweiterung der Bandbreite von Informationen, die sich zur Unterstützung von Entscheidungen nutzen lassen und schließlich zur Entdeckung neuer Muster in großen Populationen beitragen.

Automatisierung ermöglicht bedarfsgerechten Einsatz von Bildmarkern

Maschinenlernen und besonders Deep Learning sind bei der Verarbeitung und Klassifikation von Bildinformationen äußerst erfolgreich. Zu den Aufgaben, die, wenn sie von Menschen ausgeführt würden, langwierig oder in angemessenem Umfang gar nicht zu bewältigen wären, zählen die Entdeckung und Klassifikation von Zellen in Mikroskopbildern¹ oder die Identifizierung und Segmentierung von Flüssigkeit in Retinal-Imaging-Daten.² Für die biomedizinische Bildsegmentierung entwickelte Convolutional Neural Networks und verschachtelte Architekturen wie U-Nets machen sich ein einfaches, aber unerhört hilfreiches Element der visuellen Welt zunutze: Kompositionalität. Anstatt in einem Durchgang hochkomplexe Strukturen zu entdecken, bestehen sie aus Elementen, die sehr einfache Strukturen wie Ränder, Häufigkeit oder Richtung erkennen. Der Schlüssel zum Leistungsvermögen dieser simplen Detektoren liegt darin, dass sie in Schichten übereinandergestapelt sind und damit die wiederholte Neuzusammensetzung von Detektoren für einfachere Elemente zu komplexeren Konfigurationen ermöglichen. Wo Merkmalextraktoren vor einiger Zeit noch von Hand programmiert wurden, werden heute tiefe neuronale Netze dafür trainiert, Merkmalextraktoren und die entsprechenden Vorhersagemodelle in einem zu konstruieren.

¹ O. Ronneberger, P. Fischer und T. Brox, »U-Net: Convolutional Networks for Biomedical Image Segmentation«, in: N. Navab, J. Hornegger u. a. (Hg.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 3 Bde., Teil 3, Cham 2015, S. 234–241.

² T. Schlegl, S. M. Waldstein u. a., »Fully Automated Detection and Quantification of Macular Fluid in OCT Using Deep Learning«, in: *Ophthalmology*, Bd. 125 (2018), Nr. 4, S. 549–558; <https://www.ncbi.nlm.nih.gov/pubmed/29224926> (27.05.2018), DOI: 10.1016/j.ophtha.2017.10.031.

A specific field that is bound to benefit from AI is precision medicine. It aims at delivering individualized care to patients. To this end, it takes the variability of disease course and the differing responses to treatment in different individuals into account. It views a disease and the corresponding treatment not as a single category, but instead as a complex pattern of individual characteristics of the patient, the disease course, and possible different responses to a range of available treatments.

A primary challenge is to identify and exploit the patterns and relationships between these complex observations and a treatment decision. Although AI has shown promising results only at the beginning, deep learning has caused excitement by solving vision-related tasks with remarkable accuracy, but has suffered from being viewed as a “black box” that performs well, but is hard to interpret at the same time. Recently, approaches to probe deep neural networks for those features that enable prediction and form an understanding of which parts of the data are informative when modelling a relationship between observation and prediction target have emerged.

It is becoming clearer that AI will serve different roles in precision medicine, ranging from the automation of repetitive tasks that require the processing of information such as images to the expansion of the range of information we can exploit to support decisions, and finally to the discovery of new patterns in large populations.

Automation enables the use of image markers at scale

Machine learning, and in particular deep learning, is highly successful in processing and classifying image information. Tasks that would be tedious or even intractable at scale if performed by humans include the detection and classification of cells in microscopy images,¹ or the detection and segmentation of fluid in retinal imaging data.² Convolutional neural networks and more involved architectures such as U-Nets developed for biomedical image segmentation exploit a simple but immensely helpful principle of the visual world: compositionality. Instead of recognizing highly complex structures at once, they consist of elements that detect very simple structures such as edges, frequency, or direction. The key to their ability is that these simple detectors are stacked in layers on top of each other and thus enable the repeated composition of detectors for simpler entities to more complex configurations. Where feature extractors were handcrafted some time ago, we now train deep neural networks to construct feature extractors and corresponding prediction models jointly.

The benefits of transferring knowledge to new problems

Radiomics is the emerging field of exploiting patterns instead of isolated markers to guide clinical decisions. In cancer diagnosis and prognosis, the integration of multiple information points is critical when preparing decisions, and the computational prediction of future disease course or outcome based on imaging features has emerged as one of the most promising directions in improving precision medicine.³ It is by now

¹ O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in N. Navab, J. Hornegger, a. o., eds., *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 3 vols., Part III (Cham, 2015), 234–241.

² T. Schlegl, S. M. Waldstein, a. o., “Fully Automated Detection and Quantification of Macular Fluid in OCT Using Deep Learning,” *Ophthalmology*, vol. 125 (2018), no. 4, 549–558; <https://www.ncbi.nlm.nih.gov/pubmed/29224926> (May 27, 2018), DOI: 10.1016/j.ophtha.2017.10.031.

³ H. J. Aerts, E. R. Velazquez, a. o., “Decoding tumour phenotype by non-invasive imaging using a quantitative radiomics approach,” *Nature Communications*, 5, article number: 4006 (2014); DOI: 10.1038/ncomms5006.

Die Vorteile der Übertragung von Wissen auf neue Probleme

Radiomics ist ein sich entwickelndes Gebiet der Nutzung von Mustern anstelle von isolierten Markern für klinische Entscheidungen. In der Krebsdiagnose und -prognose spielt die Einbeziehung einer Vielzahl von Informationen eine wesentliche Rolle für die Vorbereitung von Entscheidungen, und die computergestützte Vorhersage zukünftiger Krankheitsverläufe oder Ergebnisse auf Grundlage von Merkmalen aus Bildgebungsverfahren hat sich zu einer der vielversprechendsten Richtungen bei der Verbesserung der Präzisionsmedizin entwickelt.³ In der Onkologie ist sie mittlerweile etabliert und hat sich sehr rasch auch auf andere Krankheiten und die Einbeziehung von mehr Informationen als nur einer einzigen Modalität zur Unterstützung von Entscheidungen ausgeweitet.

Da eine große Anzahl von Trainingsbeispielen erforderlich ist, um zu leistungsstarken Modellen zu gelangen, haben sich Ansätze zur intelligenten Nutzung verschiedener Daten entwickelt. Eine zentrale Einsicht in diesem Zusammenhang war, dass Teile von Vorhersagemodellen sich übertragen und für neue Probleme wiederverwenden lassen. Ein Erkennungsmodell für spezifische Befunde aus medizinischen Bilddaten von null an zu trainieren würde durch Datenverfügbarkeit oder Rechenzeit behindert werden. Stattdessen kann man von einem vortrainierten Modell ausgehen, das zur Verarbeitung von Bildern aus der Natur benutzt worden ist, da niedrigere Ebenen, die auf grundlegende Bildelemente reagieren, sich bei der medizinischen Aufgabe wiederverwenden lassen. Kluge Feineinstellungsstrategien fokussieren den Trainingsaufwand auf Krankheiten kodierende höhere Ebenen oder anatomiespezifische Charakteristika des Prognoseproblems.

Von schwach verbundenen Bildern und Textinformationen lernen

Eine Grundidee von KI ist überwachtes (»supervised«) Lernen, bei dem die Algorithmen während des Trainings mit Beispielen gefüttert werden, die aus Beobachtungen und erwünschtem Output wie einer großen Zahl von Mikroskopbildern und Segmentierungen verschiedener Zelltypen bestehen. In der Medizin ist das oftmals nicht machbar, da der zur Erfassung der Variabilität in der Natur erforderliche Trainingsaufwand unsere Annotationsfähigkeit übersteigt. Das hat zur Verbreitung von Ansätzen geführt, die beim sogenannten schwach überwachten Lernen schwache Verbindungen zwischen Beobachtungen zum Lernen nutzen. Beispiel dafür sind etwa Ansätze, die sich klinische Routinebilder und entsprechende Aktenaufzeichnungen zunutze machen, um solche Detektoren zu trainieren. Durch die Verbindung von Bildern und radiologischen Berichten wurden damit vielversprechende Ergebnisse erzielt.⁴ Selbst die Unterweisung von Algorithmen in der Erkennung räumlicher Konzepte aus Bilddaten und entsprechenden Beschreibungen macht Fortschritte.

Verarbeitung und Erkennung komplexer Muster zur Entdeckung von Krankheitsanzeichen, zur Abschätzung von Diagnosen oder zur Vorhersage von Ergebnissen sind die Hauptanwendungen von KI zur Lösung uns bekannter Probleme. Gleichzeitig hat die Verfügbarkeit zunehmend leistungsfähiger KI-Algorithmen zu neuen Zugängen bei der Unterstützung von Diagnose- und Therapieentscheidungen geführt. Während das in der ungeheuren Zahl be-

³ H.J. Aerts, E.R. Velazquez u. a., »Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach«, in: *Nature Communications*, 5, Artikelnr.: 4006 (2014); DOI:10.1038/ncomms5006.

⁴ J. Hofmanninger und G. Langs, »Mapping Visual Features to Semantic Profiles for Retrieval in Medical Imaging«, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, S. 457–465; DOI: 10.1109/CVPR.2015.7298643.

well established in oncology and rapidly spread to other diseases and to the integration of more information for decision support than only a single modality.

As the number of training examples necessary to obtain powerful models is large, approaches to make intelligent use of different data have emerged. A key insight was that parts of prediction models can be transferred and reused for new problems. Training a recognition model for specific findings in medical imaging data from scratch would be hampered by the availability of data or processing time. Instead, we can start from a pretrained model that was used to process natural images, since lower levels that respond to basic image primitives can be reused in the medical task. Clever fine-tuning strategies focus the energy of training on higher levels that encode disease or anatomy-specific characteristics of the prediction problem.

Learning from weakly linked images and textual information

A basic idea of AI is supervised learning when during training the algorithms are presented with examples consisting of observations and the desired output such as a large set of microscopy images and segmentations of different cell types. In medicine, this is often not feasible, since the amount of training data necessary to capture the variability in nature exceeds our ability to annotate. This has motivated the ascent of approaches that exploit weak links between observations to learn during so-called weakly supervised learning. Examples are approaches that exploit clinical routine imaging and corresponding record information to train such detectors. They have shown promising results relying on linking information in images and radiology reports.⁴ Even the teaching of algorithms to recognize spatial concepts from imaging data and corresponding descriptions is advancing.

The processing and recognition of complex patterns to detect traces of disease, to estimate diagnosis, or to predict outcome are prime uses of AI to solve problems we know. At the same time, the availability of increasingly powerful AI algorithms has led to new approaches supporting diagnosis and treatment decisions. While the knowledge encoded in the immense body of cases that have been diagnosed already represents a unique source of understanding of disease paths, it has been almost impossible to access this knowledge at the point of care. Recent advances make the matching of individual cases to this body possible. Here, instead of automating repetitive tasks, AI supports diagnosis and treatment decisions by drawing the relevant information and evidence from a large and heterogeneous population of patients. This population reflects realistic paths through disease and treatment and is highly informative when compared to a case at hand.

Finally, machine learning offers a chance to expand our vocabulary of diagnostic categories capturing the interaction of patient, disease, and observable findings. Diagnostic categories and the findings we use for their determination in individual patients are a result of research and the necessity to communicate them effectively during medical training. They

⁴ J. Hofmanninger and G. Langs, "Mapping Visual Features to Semantic Profiles for Retrieval in Medical Imaging," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, 457–465; DOI: 10.1109/CVPR.2015.7298643.

reits diagnostizierter Fälle kodierte Wissen eine einzigartige Quelle zum Verständnis von Krankheitsverläufen darstellt, war es bisher fast unmöglich, am Behandlungsort auf dieses Wissen zuzugreifen. Aktuelle Fortschritte machen eine Abgleichung individueller Fälle mit diesem Korpus möglich. Anstatt sich wiederholende Aufgaben zu automatisieren, unterstützt KI hier Diagnose- und Therapieentscheidungen, indem sie aus einer großen und heterogenen Patientenpopulation relevante Informationen und Beweise herausfiltert. Diese Population spiegelt realistische Krankheits- und Therapieverläufe wider und ermöglicht hochinformativ Vergleiche mit dem jeweils vorliegenden Fall.

Schließlich bietet Maschinenlernen die Chance, durch die Erfassung der Interaktion von Patient, Krankheit und feststellbaren Befunden das Vokabular diagnostischer Kategorien zu erweitern. Diagnostische Kategorien und die Befunde, die wir verwenden, um sie bei individuellen Patienten festzustellen, sind das Resultat von Forschung und der Notwendigkeit, sie im Verlauf der medizinischen Ausbildung wirksam zu vermitteln. Sie sind ein Mittel zum Zweck: eine optimale, auf exakten Prognosen der Therapiereaktion basierende Therapieentscheidung. Aufgrund seiner Eignung, verteilte komplexe Auffindungsmuster und rechnerische Gruppenidentifikation auf extrem große Patientenpopulationen anzuwenden, beginnt sich das Maschinenlernen damit auf diese Kategorien auszuwirken. Ergebnisse zeigen, dass unüberwachtes Maschinenlernen in Routinebilddaten, die mit klinisch relevanten Problemen in Verbindung stehen, Phänotypen aufzufinden vermag.⁵ Die Erweiterung unseres Befundvokabulars wird also durch die Entdeckung von Anomalien ermöglicht, die nicht unter bestehende Marker Kategorien fallen.

Wo stehen wir heute?

KI und vor allem Maschinenlernen verheißen einen veränderten Ansatz in der Präzisionsmedizin. Sie laufen darauf hinaus, bei individuellen Patienten und ihren Krankheitsgeschichten die Erkennung und Verwendung von Mustern zu ermöglichen, die eine optimale Behandlung bieten. Die Rolle von KI wird sich von der Automatisierung einfacher Aufgaben bis zur hochentwickelten Interaktion zur Unterstützung der Entscheidungsprozesse klinischer Experten erstrecken. Wir stehen immer noch erst am Anfang dieser Reise. Zu den Herausforderungen zählen die Verwendung häufig unstrukturierter und heterogener Daten für das Trainieren von Vorhersagemodellen sowie die Subtilität der Anzeichen verschiedener Varianten einer Krankheit, welche substanzielle Auswirkungen auf die optimale Behandlungsentscheidung haben, dabei aber schwer von natürlicher Variabilität zu unterscheiden sind. Schließlich gilt es zu durchschauen, wie sich das Zusammenwirken und die wechselseitige Ergänzung der Beiträge von klinischen Fachleuten und KI zu einer individualisierten Therapie gestalten werden. Eines ist sicher: Der Weg dahin führt über interdisziplinäre Zusammenarbeit von Medizin und Computerwissenschaft. ✕

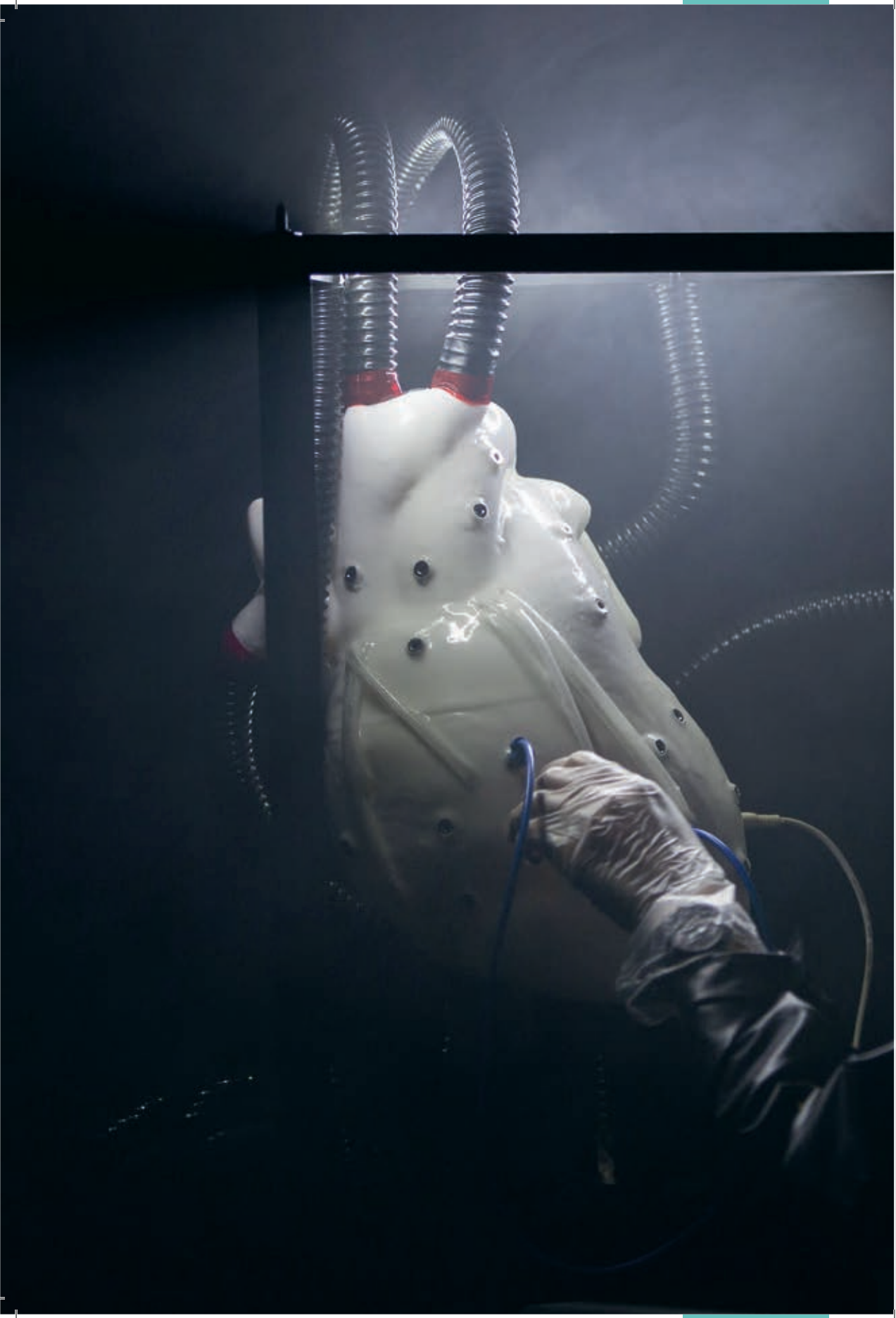
⁵ J. Hofmanninger, M. Krenn u. a., »Unsupervised Identification of Clinically Relevant Clusters in Routine Imaging Data«, in: S. Ourselin, L. Joskowicz u. a. (Hg.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, 3 Bde., Teil I, Cham 2016, S. 192 – 200; doi: https://doi.org/10.1007/978-3-319-46720-7_23.

are means to an end: an optimal treatment decision based on accurate predictions of treatment response. Thus, machine learning is starting to affect these categories, due to its ability to use distributed complex finding patterns and the computational identification of groups in extremely large patient populations. Results show that unsupervised machine learning can find phenotypes in routine imaging data that are linked to clinically relevant problems.⁵ The expansion of our findings vocabulary is enabled by the detection of anomalies not falling into existing categories of markers.

⁵ J. Hofmanninger, M. Krenn, a. o., "Unsupervised Identification of Clinically Relevant Clusters in Routine Imaging Data," in S. Ourselin, L. Joskowicz, a. o., eds., *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016*, 3 vols., Part I (Cham, 2016), 192–200; doi: https://doi.org/10.1007/978-3-319-46720-7_23.

Where are we right now?

AI and in particular machine learning are promising to change our approach toward precision medicine. They are bound to enable the identification and use of patterns in individual patients and their disease history to offer optimal treatment. The role of AI will range from the automation of simple tasks to a sophisticated interaction supporting clinical experts in their decision process. We are still at the beginning of this journey. Challenges include the use of often unstructured and heterogeneous clinical data for the training of prediction models and the subtle traces of different variants of disease that have substantial impact on the optimal treatment decision but are hard to discern from natural variability. Finally, we will have to learn how the interaction and complementary contribution to individualized treatment of clinical professionals and AI will emerge. One thing is certain: the path to this runs through interdisciplinary work among medicine and computer science. ✕



W
Y

»
m
3-
la
S
m
d
zi
V
w
vi
a
s
e
v
b
z
M
u

“
s
c
a
la
a
w
p
s
a
fr
s
is
sl
a
st

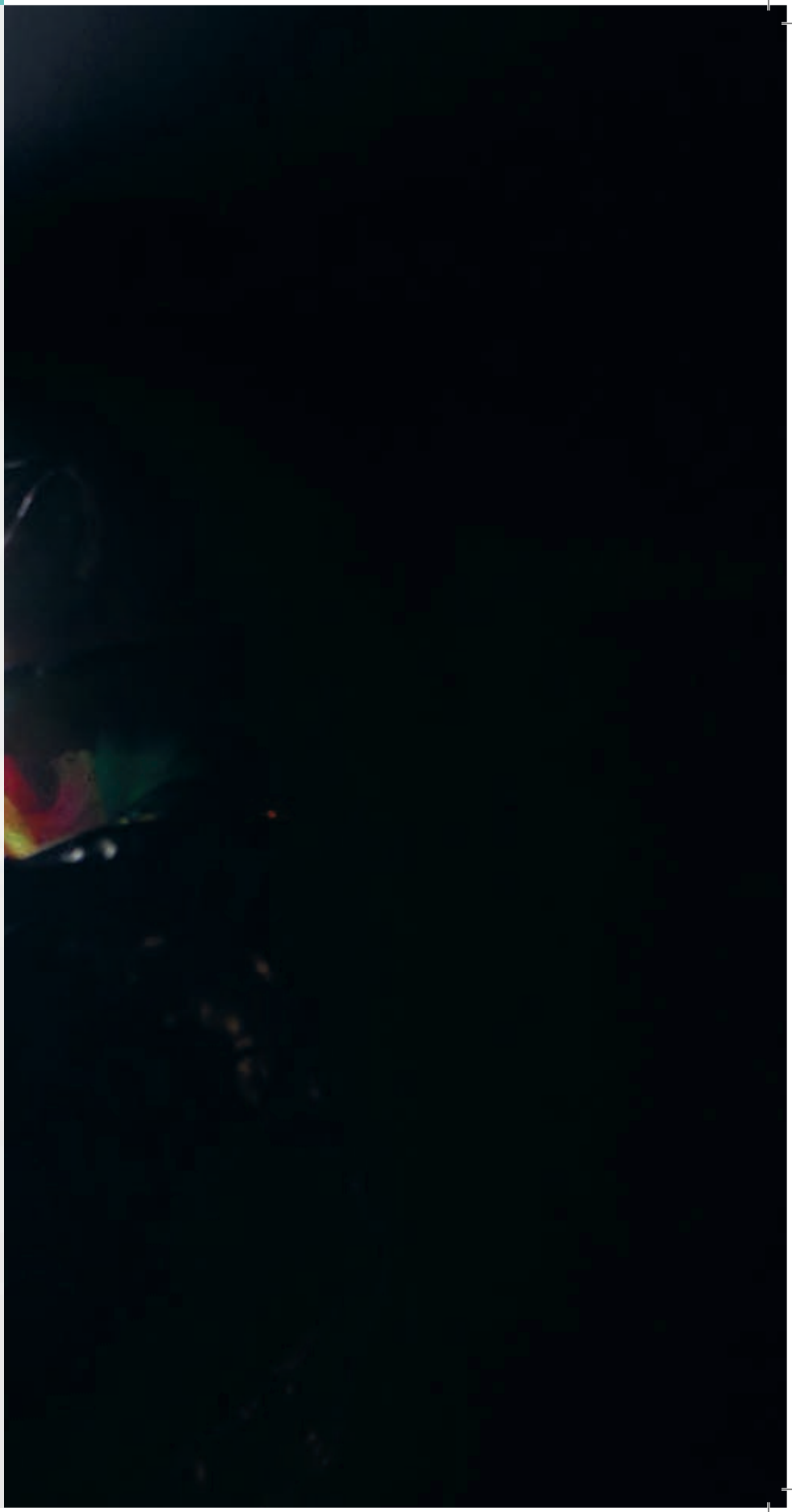
—
©
F

Whose scalpel **Yen Tzu Chang (tw)**

»Whose scalpel« ist eine mit einer visuellen und 3-D-gedruckten Installation kombinierte Soundperformance, die mithilfe eines Anwendungsrahmens für medizinische bildgebende Verfahren realisiert wurde. Das Projekt, das verschiedene Methoden aus Kunst und Wissenschaft vereint, stellt ein Zukunftsszenario vor und wirft Fragen bezüglich der Beziehung zwischen Mensch und Maschine in der Herzchirurgie auf.

“Whose scalpel” is a sound performance combined with a visual and 3D-printed installation, realized with an application framework for medical-image processing. Mixing several methods from art and science, it confronts us with a future scenario and highlights issues of the relationship between man and machine in heart surgery.

© Ars Electronica/
Florian Voggeneder



Künstliche Intelligenz als dienstbarer Helfer

Es wird in absehbarer Zukunft kaum einen Dienstleistungsbereich geben, der nicht von künstlicher Intelligenz beeinflusst oder sogar revolutioniert werden wird. – Ein kurzer Überblick von A wie Agrarwirtschaft bis V wie Verkehr.

Alexa, Echo, Siri & Co sind zwar noch nicht lange auf dem Markt, doch ihre Erfolge sind, zumindest bisher, überwältigend. Man sagt etwas, und schon wird zum Beispiel das Licht gedimmt, die Heizung gedrosselt, ein Song eingespielt, eine neue Packung Milch bestellt oder, oder, oder. Diese digitalen Assistenten arbeiten mit Algorithmen zur Erkennung und Verarbeitung natürlicher Sprachen (natural language processing), die durch neuronale Netzwerke in jüngster Zeit enorme Fortschritte gemacht haben. Die Spracherkennung ist die erste große Disruption, die künstliche Intelligenz (KI) hervorgebracht hat. Jüngst sorgte etwa Google mit einem Assistenten für Aufsehen, der Aufgaben wie die Reservierung eines Tisches im Restaurant oder eines Termins beim Friseur erledigte – wobei dem Gegenüber nicht auffiel, dass er mit einem Computer sprach.

Sprachsteuerung ist in den Augen vieler Technologiegurus auch drauf und dran, die Art und Weise, wie wir Computer bedienen, zu ergänzen, wenn nicht gar zu ersetzen. Während einem derzeit die Hände bei der Schreibtischarbeit im wahrsten Sinn des Wortes gebunden sind – und zwar an Tastatur und Maus –, könnte die Bedienung per Spracheingabe den Berufsalltag der meisten Menschen völlig verändern. Eine weitere denkbare Art der Steuerung von Computersystemen, nämlich durch das direkte Erkennen von Gedankenimpulsen, steckt im Vergleich dazu noch völlig in den Kinderschuhen.

Innovative Verkehrssysteme

Praktisch alle Experten meinen, dass der Verkehrssektor der nächste große Bereich sein wird, in dem KI für tiefgreifende Revolutionen sorgen wird. Das betrifft zum einen die Entwicklung von autonomen Fahrzeugen. In diesem Bereich läuft derzeit ein Match zwischen den großen etablierten Herstellern und den vielen Newcomern in der Branche darum, wer als Erstes mit einem in der Praxis funktionstauglichen Auto auf den Markt kommen wird. Die jüngsten Erfahrungen, wie etwa das erste Todesopfer bei einer Testfahrt, haben gezeigt, dass noch viele Probleme gelöst werden müssen. Auf dem Weg dorthin werden die Fahrzeuge allerdings mit immer mehr Assistenzsystemen ausgestattet – von Notbremssystemen über Spurhalteassistenten bis hin zu Einparkautomaten. Dank Car-to-Car-Kommunikation wird es in naher Zukunft möglich werden, dass man quasi um Kurven herumschauen kann.

Artificial Intelligence as a Subservient Aide

**In the near future there will hardly be a service sector left that will not have been influenced or even revolutionized by artificial intelligence.—
A brief survey from A for agriculture to W for waste management.**

Although Alexa, Echo, Siri, and their ilk have not been on the market for a very long time, their success has been stunning, at least until now. You say something, and in next to no time the light will be dimmed, the heating will be turned down, a particular song will be played, a carton of milk will be ordered, and so on and so forth. These digital assistants work with algorithms for speech recognition and natural language processing that have improved enormously lately, thanks to neural networks. Speech recognition is the first major disruption artificial intelligence (AI) has brought forth. Recently, Google caused a sensation with an assistant capable of carrying out such tasks as reserving a table at a restaurant or making an appointment at the hairdresser's—with the persons on the other end of the line not realizing that they were conversing with a computer.

Many technology gurus believe that voice control is about to complement if not replace the way of how we operate computers. Whereas our hands are still literally tied—to the keyboard and mouse—when we perform desk-based work, operating devices controlled via voice command might profoundly change office routines for most people. Compared to this, another conceivable method of controlling computer systems, namely through their direct recognition of mental impulses, is still in its infancy.

Innovative traffic and transportation systems

Practically all experts agree that transportation will be the next major sector in which AI will bring about fundamental revolutions. For one thing, this concerns the development of autonomous vehicles. There is currently a match going on between the big established automobile manufacturers and the industry's many newcomers about who will be the first to market a car that is fully functional in practice. The most recent experiences, including the first death during a test drive, have shown that many problems still await solution. In the meantime, however, vehicles are being furnished with more and more assistance systems—from autonomous emergency braking and lane departure warning systems to automatic parking systems. In the near future it will be possible to virtually look around bends thanks to vehicle-to-vehicle communications.

Zum anderen aber – und das ist weniger offensichtlich – wird KI das gesamte Verkehrsgeschehen grundlegend verändern. Sichtbar sind etwa bereits Informationssysteme über Anschlüsse von öffentlichen Verkehrsmitteln oder wirklich intelligente Routenplaner, die auch das Verkehrsgeschehen und das Wetter berücksichtigen. Auch die Verkehrsinfrastruktur wird zunehmend intelligent: Sensoren erfassen das gesamte Verkehrssystem, verknüpfen diese Daten mit vielen anderen Parametern und liefern so Inputs für die optimale Steuerung des Verkehrsflusses. Durch zunehmende Verknüpfung von Infrastruktur und Fahrzeugen entstehen kooperative Verkehrssysteme. Zudem gibt es laufend Verbesserungen beim Routing etwa von Car-Sharing- oder neuen Taxi-Systemen. Ihr volles Potenzial werden solche neuen Dienstleistungen dann wohl in Verbindung mit autonomen Fahrzeugen ausspielen.

KI in der Finanzwelt

Auch im Finanzsektor gibt es vielfältige Einsatzszenarien für KI. Automatische Handelssysteme, E-Commerce oder Onlinebanking sind ja längst keine Neuigkeit mehr, aber je intelligenter diese Systeme werden, desto mehr werden sie an Bedeutung gewinnen. KI kann aus der Analyse vergangener Transaktionen Betrugsmuster in Transaktionsgeschäften erkennen und möglicherweise Fehler automatisch korrigieren. Viel Kritik gab es in jüngster Vergangenheit an neuen Methoden zur Bewertung von Kunden hinsichtlich ihrer Kreditwürdigkeit oder Versicherbarkeit. Manche glauben, dass KI vor allem in Verbindung mit der Blockchain-Technologie den gesamten Sektor verändern wird: Zentrale Vermittler für Transaktionen sind dabei nicht mehr notwendig, und mithilfe von »Smart Contracts« werden viele Routinetätigkeiten obsolet. Auch im Kleinen verändert die Technologie zunehmend den Umgang mit Geld: So wurde kürzlich in London ein System vorgestellt, mit dem man Straßenmusikern bargeldlos eine Spende geben kann. Das funktioniert, ähnlich wie bei NFC-Kassen in Supermärkten, mithilfe von Karten, Smart Watches oder Smartphones mit entsprechenden Funkchips.

Juristerei und Kommunikation

Ein Gebiet, das sich sehr gut für die Automation durch KI eignet, sind Rechtsdienstleistungen. Kaum ein anderer Bereich baut dermaßen stark auf präzise definierte Regeln (Gesetze etc.) in einer stark formalisierten Sprache auf. Dank der Fülle an schriftlich dokumentierten Fällen und Urteilen stehen für maschinelle Lernsysteme reichlich Trainingsdaten zur Verfügung. Wie groß die Potenziale sind, zeigte vor einigen Jahren der vergleichsweise simple Legal-Bot »DoNotPay«, der von einem 19-jährigen Studenten der Stanford University programmiert wurde. Dieses Programm kann automatisiert Einsprüche gegen Parkstrafen verfassen, indem der Bot beim Betroffenen die konkreten Umstände des – angeblichen – Delikts abfragt und am Ende eine juristisch einwandfreie Berufung auswirft. Bei komplexeren Rechtsmaterien und höheren Strafen bzw. möglichen Schäden vertrauen Juristen freilich noch nicht zu 100 Prozent den KI-Helfern. Doch in vielen Anwaltskanzleien sind bereits Systeme zur Dokumentenrecherche, zum Durchsuchen von Präzedenzfällen oder zur Suche

On the other hand, albeit less obviously, AI will essentially transform transportation as a whole. First visible signs include systems providing information on public transport connections or truly intelligent route planners taking into account the current traffic situation and weather conditions. Traffic infrastructure too is becoming increasingly intelligent: sensors tracking the traffic system in its entirety combine their data with numerous additional parameters and, as a result, deliver input for an optimal control of traffic flow. This enhanced linkage between infrastructure and vehicles leads to the emergence of cooperative traffic systems. Furthermore, the routing functions of car-sharing or new taxi systems undergo constant improvement. These new services will probably be playing to their full potential in conjunction with autonomous vehicles.

AI in the world of finance

The financial sector similarly has a broad spectrum of application scenarios for AI in store. Automated commercial systems, e-commerce, and online banking have long ceased to be rated as novelties, but the more intelligent these systems become, the more significant their future role. AI can detect fraud patterns in business transactions from the analysis of past transactions and possibly correct mistakes automatically. Recently, new methods of assessing clients as to their financial soundness or insurability have provoked severe criticism. Some believe that AI will change the whole sector, above all when teamed up with the blockchain technology: central agents orchestrating transactions will be made redundant, and thanks to smart contracts many routine jobs will become obsolete. The new technology increasingly changes the way money is handled, also on a small scale: lately a system was introduced in London that allows people to make cashless donations to street performers. This functions similarly to NFC technology used in supermarkets, with the aid of cards, smart watches, or smartphones equipped with suitable radio chips.

Jurisprudence and communication

Legal services ideally lend themselves to automation through AI. Hardly any other field relies on such precisely defined rules (laws, etc.) expressed in such heavily formalized language. Thanks to the abundance of cases and judgments documented in writing, ample training data is available for machine learning systems. Several years ago, “DoNotPay,” a comparatively simple legal bot programmed by a nineteen-year-old student from Stanford University, demonstrated the huge potentials it held. Its program is capable of automatically disputing parking tickets: having inquired the concrete details of the—alleged—offender, the bot will eventually come up with a legally flawless notice of appeal. In more complex legal matters or when higher fines or possible damage is involved, attorneys naturally do not yet completely rely on AI assistants. Nevertheless, many law offices already employ automated systems for

von Lücken in Verträgen im Einsatz. Daraus ergibt sich in der Praxis eine starke Reduktion der dann noch »händisch« von erfahrenen Juristen zu durchforschenden Dokumente.

KI wird auch die Kommunikation und den Journalismus verändern. In der Werbewirtschaft werden vor allem im Onlinebereich schon heute Werbepläne von intelligenten Algorithmen gesteuert, welche die Botschaften möglichst zielgerichtet bei den geeigneten Konsumenten ankommen lassen. Manche Medienunternehmen setzen bereits »Roboterjournalisten« ein, die bei genau definierten Sachverhalten selbsttätig aktuelle Ereignisse aus dem Newsflow identifizieren und zu lesbaren Texten zusammenbauen. Das funktioniert etwa bei Finanznachrichten oder bei Berichten über diverse Sportarten bereits tadellos. Allerdings sind zum Beispiel bei Analysen, Hintergrundberichten und erst recht bei Kommentaren unverändert der Sachverstand, die Erfahrung und das Gespür leibhaftiger Journalisten gefragt.

Sicherheit

Große Potenziale, aber gleichzeitig wohl auch die größten Ängste vor Missbrauch, bestehen im Bereich Sicherheit und Überwachung. Das beginnt beim Aufbau von Kameranetzwerken, mit denen einzelne Individuen identifiziert und auf all ihren Wegen verfolgt werden können, und reicht bis hin zur Verknüpfung komplexer Datenbanken – etwas, was schon vor Jahren unter dem Schlagwort »Rasterfahndung« heiß diskutiert wurde, aber mit damaligen Technologien vergleichsweise harmlos war, wenn man die rasante Weiterentwicklung von KI-Methoden seit jener Zeit in Betracht zieht. Methoden wie »Predictive Policing« (vorausschauende Polizeiarbeit), ein Verfahren, mit dem man die Wahrscheinlichkeit zukünftiger Straftaten berechnen kann, lassen sowohl Datenschützern als auch Bürgerrechtlern die Grausbirnen aufsteigen. Als warnende Beispiele gingen in jüngster Zeit Entwicklungen in China durch die Medien: Dort werden Milliarden von Einträgen in Datenbanken oder Bilder der zahlreichen Überwachungskameras zu einem sogenannten »Sozialkreditsystem« verknüpft, das etwas über das Wohlverhalten jedes einzelnen Bürgers aussagen soll. Es gibt Abzüge, wenn jemand beispielsweise bei Rot über die Kreuzung geht oder gegen Zensurvorschriften verstößt, und Gutpunkte, wenn jemand z. B. kommunistische Lieder singt. Der Punktestand entscheidet schließlich darüber, welche Leistungen ein Individuum nutzen darf – ob es etwa als kreditwürdig gilt oder als Passagier für einen Schnellzug in Frage kommt. Solche Systeme nehmen in China derzeit überhand: Kürzlich wurde von einer Schule in Hangzhou berichtet, in der eine Software eingesetzt wird, die anhand von Kamerabildern die Gesichter der Schüler analysiert und deren Aufmerksamkeitsniveau ermittelt – offiziell zu dem Zweck, dass Lehrende ihren Unterricht optimieren können.

document research, going through previous court decisions, or identifying contractual gaps. In practice, this will result in a substantial reduction of “manual” document browsing performed by experienced lawyers.

AI is also going to transform communication and journalism. In today’s advertising industry—mainly in the online sector—advertising schedules are already coordinated by intelligent algorithms directing messages in the most efficient way imaginable at potentially interested consumers. Some media companies already count on “robot journalists” automatically identifying current events in the information flow for accurately defined contexts and combining them to create legible texts. This works perfectly for financial news or reports on various sporting events, for example. However, when it comes to analyses, background reviews, and, above all, editorial comments, the expertise, experience, and instinct of living journalists continue to be vital ingredients.

Security

The field of security and surveillance harbors huge potentials and at the same time causes the greatest anxieties as to data abuse. This ranges from the installation of camera networks able to identify individuals and observe them wherever they go to the integration of complex databases—an issue fiercely discussed years ago in the context of dragnet investigation and police profiling, which, given the technologies of those days, seemed still rather harmless then when compared to the rapid development of AI methods that has taken place in the meantime. Such methods as predictive policing, a technique that makes it possible to compute the probability of future crimes, have alerted privacy advocates and civil rights activists.

Most recently, developments in China have been doing the rounds in the media as warning examples: billions of database entries or images from innumerable surveillance cameras are fed into a so-called “Social Credit System,” which is meant to rate the reputation and integrity of each individual citizen. For example, people crossing the street on a red light or violating censorship regulations receive minus points, while those chanting communist songs earn bonus points. The score finally decides about the services for which an individual is eligible—for example, if he or she is considered creditworthy or allowed to ride express trains. Such systems are currently getting out of hand in China: there are recent reports about a school in Hangzhou that uses software to analyze the faces of students by means of camera images and determine their level of attention—for the official purpose of optimizing teaching methods.

Von der Abfall- zur Agrarwirtschaft

Trotz aller – vielleicht auch berechtigten – Dystopien wird es kaum eine Branche geben, in der KI keine Rolle spielen und Nutzen stiften wird. Hier in aller Kürze nur drei weitere Beispiele für Dienstleistungen in Bereichen, die sehr weit voneinander entfernt sind. Da wären etwa die Abfall- und die Bauwirtschaft, die große Stoffströme möglichst optimal transportieren müssen. Dafür werden derzeit intelligente Logistiksysteme aufgebaut, die alle relevanten Daten miteinander verknüpfen. Diese Modelle sind auch die unabdingbare Voraussetzung dafür, die immer strenger werdenden Recyclingvorgaben zu vertretbaren Kosten zu erfüllen.

Ein zweites Beispiel ist die Kunstgeschichte: In Archiven, Hof- und Klosterbibliotheken stauben Abermillionen von unaufgearbeiteten Dokumenten vor sich hin. Hier können digitale Technologien wie automatische Bilderkennung oder intelligente Datenbanken eine unschätzbare Hilfe bei der Dokumentation oder gezielter Suche nach Antworten auf Forschungsfragen bieten.

Und schließlich setzt auch die Landwirtschaft zur Zeit voll auf Digitalisierung, um zum einen Ressourcen zu sparen und die Ernten zu optimieren und zum anderen die Auswirkungen auf die Umwelt zu minimieren. Ein Beispiel aus China zeigt, welch gigantischer Aufwand hier schon betrieben wird: Aufgebaut wurde ein riesiges Netzwerk von 1200 Agrarexperten, 65.000 lokalen Beamten, 140.000 Industrievertretern und 21 Millionen Bauern, die über Jahre alle möglichen Informationen über Bewässerung, Pflanzendichte oder Saattiefe gesammelt und mit der Produktivität in Beziehung gesetzt haben. Dadurch wurde es möglich, die Ernten um gut ein Zehntel auszuweiten und gleichzeitig den Einsatz von Dünger um ein Sechstel zu reduzieren. ✕

From waste management to agriculture

Despite all the dystopias, as justified as they may be, there will be hardly a sphere where AI will not play a crucial role and which will not benefit from it. This is to briefly introduce three further examples of services that could not be further apart. In the waste and building industries, for instance, huge amounts of materials need to be transported as efficiently as possible. For this purpose, intelligent logistic systems are being installed these days to pool all the relevant data. These models are also categorical prerequisites for being able to meet recycling requirements, which are getting stricter and stricter, at reasonable costs.

Another example is art history: in archives and court and monastery libraries, millions upon millions of unprocessed documents are gathering dust. Digital technologies like automated image analysis or intelligent databases offer invaluable help in compiling documentations or conducting targeted searches to find answers to research questions.

Finally, agriculture has come to fully rely on digitization nowadays—to save resources and optimize crop yields on the one hand and to minimize negative impacts on the environment on the other. An example from China illustrates the gigantic effort that has already been put in here: a huge network consisting of 1,200 agricultural experts, 65,000 local officers, 140,000 industry representatives, and 21 million farmers has been installed to compile all kinds of information on irrigation, crop density, and sowing depth over the years and relate them to productivity. In this way, it has become possible to increase crop yields by at least ten percent while reducing the use of fertilizers by one sixth. ✕

Helga Nowotny/Stefan Thurner

Wie im Jahr 2038 Entscheidungen gefällt werden



Entscheidungsprozesse der 2010er-Jahre haben zu oft unbedachten sozialen Folgen geführt. 2038 lässt sich dank der Rechenkapazität von HUB-OLDT 4.0 und der Lernfähigkeit seiner Algorithmen das gesellschaftliche Leben eines ganzen Landes in einem erweiterten Wirklichkeitsraum abbilden. Vor einer Entscheidung kann jedes erdenkliche Szenario virtuell durchgespielt, auf mögliche Auswirkungen getestet und einer breiten Öffentlichkeit kommuniziert werden. Eine Vision.

Wir schreiben das Jahr 2038 – und feiern heute eine Errungenschaft des frühen 21. Jahrhunderts, deren Geschichte 2018 begann. Eine längerfristige Finanzierung durch die Österreichische Nationalstiftung ermöglichte es dem Complexity Science Hub Vienna, einen Simulator als Prototyp herzustellen, der heute Hunderte Millionen von Personen als Avatare abbildet und ganze Bevölkerungen virtuell simulieren kann. Anfänglich im Keller des Palais Strozzi entwickelt, steht der zentrale Rechner heute, 2038, in einem aufgelassenen Stollen der Salinen Austria, der zu den bestgesicherten Rechnerstandorten Europas zählt. Biometrisches Einloggen erlaubt einen weltweiten Zugriff auf den Simulator – nicht nur von Regierungen, sondern von weiten Teilen der Bevölkerung.

HUB-OLDT 4.0, so der Name des Simulators, gibt faktenbasierte Antworten auf Fragen von gesellschaftspolitischer Dimension. Er berechnet, ausgehend von sämtlichen Daten, die in einer Gesellschaft zur Verfügung stehen, die Konsequenzen von politischen, infrastrukturellen, finanziellen und ökonomischen Entscheidungen, lange bevor diese in der Wirklichkeit getroffen werden. Jedes erdenkliche Szenario kann virtuell durchgespielt werden – von der Etablierung einer Fußgängerzone an der Südosttangente bis zur Planung des optimalen Ressourceneinsatzes in der Republik im Hinblick auf eine maximale Innovationsrate bei gleichzeitiger sozialer Gerechtigkeit.


Größere Zusammenhänge werden sichtbar

Der Simulator wird von einer immensen Anzahl von Daten gespeist, die es ermöglichen, den zukünftigen Verlauf komplexer sozioökonomischer Prozesse zu extrapolieren. Statt Daten und Fakten isoliert zu betrachten und in Datensilos zu sammeln, macht HUB-OLDT 4.0 größere Zusammenhänge sichtbar. Die Berechnung erfolgt durch Algorithmen, die sich durch neu hinzukommende Daten ständig erneuern und anpassen – »machine learning« at its best: Die Maschine lernt durch die programmierten Regeln und eingespeis-

Helga Nowotny ist eine international anerkannte Wissenschaftsforscherin und Professorin emerita der ETH Zürich. Sie war Gründungsmitglied und Vizepräsidentin des 2007 etablierten Europäischen Forschungsrats und von 2010 bis 2013 dessen Präsidentin. Sie ist Vorsitzende des ERA Council Forum Austria und Mitglied des Rates für Forschung und Technologieentwicklung.



Stefan Thurner ist Physiker und Komplexitätsforscher. Seit 2009 ist er Professor für die Wissenschaft komplexer Systeme an der Medizinischen Universität Wien. Er ist externer Professor am Santa Fe Institute und seit 2015 Präsident des Complexity Science Hub Vienna. Vom Klub der Bildungs- und Wissenschaftsjournalisten wurde er als österreichischer Wissenschaftler des Jahres 2017 ausgezeichnet.




Helga Nowotny is an internationally recognized theorist of science and professor emerita of the ETH Zurich. She was a founding member and Vice-President of the European Research Council and its President from 2010 to 2013. She is chair of the ERA Council Forum Austria and member of the Council for Research and Technology Development.

Helga Nowotny/Stefan Thurner

How Decisions Will Be Made in 2038

Decision-making processes of the 2010s have often had unforeseen social consequences. In 2038, thanks to the computing capacity of HUB-OLDT 4.0 and the learning capacity of its algorithms, the social life of an entire country can be represented in an extended reality space. Prior to any decision, any conceivable scenario can be simulated, tested with regard to its potential impact, and communicated to a wider public. A vision.

We write the year 2038—celebrating today an achievement of the early twenty-first century whose history began in 2018. Long-term funding by the Austrian National Foundation enabled the Complexity Science Hub Vienna to build a prototype simulator that is able today to represent hundreds of millions of people as avatars and to virtually simulate entire populations. Originally developed in the basement of Palais Strozzi in Vienna, the central mainframe computer is installed today, in 2038, deep inside a disused salt mine formerly operated by Salinen Austria, one of the best-protected computer locations in Europe. A biometric login system permits worldwide access to the simulator—not only by governments but also by large segments of the general population.



Stefan Thurner is a physicist and complexity researcher. He holds a full professorship for the Science of Complex Systems at the Medical University of Vienna since 2009. He is external professor at the Santa Fe Institute and, since 2015, president of the Complexity Science Hub Vienna. He was awarded the title Austrian “Scientist of the Year 2017” by the Club of Education and Science Journalists.

HUB-OLDT 4.0, which is the name of the simulator, provides fact-based answers to questions of a sociopolitical dimension. Based on all data available in a society, it computes the consequences of political, infrastructural, financial, and economic decisions long before they are made in reality. Any conceivable scenario can be virtually simulated—from establishing a pedestrian zone alongside an urban ring highway to the planning of optimum allocation of resources in the republic so as to facilitate a maximum rate of innovation while simultaneously securing social justice.

Getting the bigger picture

The simulator is fed with an immense amount of data, which enable extrapolation of the future trajectories of complex socioeconomic processes. Instead of isolated fact and data analysis and collection in data silos, HUB-OLDT 4.0 brings out the wider contexts. The computation is done by algorithms that keep renewing themselves and adapting through new data input—“machine learning” at its best: the machine keeps learning continuously from programmed rules and fed-in data. In the

ten Daten ständig dazu. Durch die Simulation können real verlaufende Prozesse dynamisch visualisiert werden, sodass sinnvolle und den Menschen verständliche Antworten auf die an das Modell gestellten Fragen erfolgen.

Der Simulator ist auf Fragestellungen ausgerichtet, die real ablaufende Prozesse aus den Bereichen Wirtschaft, Gesundheit, Sicherheit, Bildung, Forschung und Innovation abbilden. So kann beispielsweise gezeigt werden, welche unbeabsichtigten Folgen entstehen, wenn etwa eine Steuer auf Produktionsmittel in der »Industrie 19.0« eingeführt würde. Hypothesen, Ideen, geplante Maßnahmen, aber auch Gedankenexperimente können auf diese Weise virtuell auf ihre möglichen Auswirkungen getestet und einer breiten Öffentlichkeit vollkommen transparent kommuniziert werden. Jeder neue Vorschlag kann sofort simuliert werden, wodurch es zu drastisch beschleunigten und wahrhaft evolutionären Planungsprozessen kommt, die wir heute als selbstverständlich voraussetzen. Heute wird keine Entscheidung mehr getroffen, die nicht vorher virtuell simuliert wurde.

In den antiquierten Entscheidungsprozessen der 2010er-Jahre sind oft unbedachte Folgen in der sozialen Wirklichkeit aufgetreten, die entweder irreversibel oder mit erheblichen und unnützen Kosten verbunden waren. Dank der Rechenkapazität von HUB-OLDT 4.0 und der Lernfähigkeit seiner Algorithmen lässt sich heute das gesellschaftliche Leben eines ganzen Landes in einem erweiterten Wirklichkeitsraum abbilden.

Die Anfänge 2018 waren bescheiden: HUB-OLDT 4.0 wurde am Hub zunächst als Forschungstool entwickelt. Innerhalb kürzester Zeit entstand eine weltweite Nachfrage nach seiner einmaligen, transparenten und anonymisierten Datenzusammenführung, die zwar echte Daten verwendet, diese aber in anonymen virtuellen Avataren so abbildet, dass die Daten keiner Person mehr zugeordnet werden können. Das war der Schlüssel, um individuelle Daten zu verarbeiten, ohne die Persönlichkeitsrechte und die Privatsphäre anzutasten. Weit über die Scientific Community hinaus zeigten Venture-Kapitalisten, Regierungen und große Konzerne brennendes Interesse an der Nutzung.

Open Simulation – Was HUB-OLDT 4.0 alles kann

Ziel war es von Anfang an, den Zugang zu den erstaunlichen Leistungen des Simulators nicht nur einem kleinen Kreis von Eingeweihten vorzubehalten. Dem Beispiel der frühen 2000er-Jahre von Open Access und Open Innovation folgend, sollte seine Simulationsexperimente nicht nur für die Scientific Community und für internationale Konzerne verfügbar sein, sondern sich für die Gesellschaft öffnen: Open Simulation (OS) war geboren. Es war ein logischer nächster Schritt in der Annäherung von Wissenschaft und Gesellschaft. Wo damals Google, die NSA, Facebook, Amazon und andere Konzerne in heute unvorstellbarer Weise mit personenbezogenen Daten den Markt dominierten, bietet HUB-OLDT 4.0 die erfolgreiche Alternative. Wo damals Google entschied, was wir zu denken, und Amazon, was wir zu kaufen haben, steht den Menschen 2038 eine wissenschaftlich und gesellschaftspolitisch sinnvolle Grundlage zur Verfügung. Ihre Entscheidungen und Handlungen erhalten einen neuen Raum – den der Simulation –, in dem die demokratische Beteiligung vollkommen anders und freier praktiziert wird als vor zwanzig Jahren.

simulation, real-life processes can be dynamically visualized so that meaningful and humanly intelligible answers are given to questions asked of the model.

The simulator is laid out for queries addressing ongoing real-life processes in the areas of economy, public health, security, education, research and innovation. So it can, for example, be demonstrated what the unintended effects would be if, for instance, a tax would be levied on means of production in the “Industry 19.0.” Hypotheses, ideas, planned measures, but also thought experiments can thus be tested for their possible impact and communicated in a completely transparent manner to a wide public. Every new suggestion can be simulated at once, which facilitates dramatically accelerated and truly evolutionary planning processes as we take for granted today. Today, no decision is made without prior virtual simulation.

In the antiquated decision-making processes of the 2010s, unconsidered consequences frequently made themselves felt in social reality that either were irreversible or entailed substantial and useless costs. Today, thanks to the computing capacity of HUB-OLDT 4.0 and the learning capacity of its algorithms, the social life of an entire country can be reproduced in an extended reality space.

The beginnings were humble: HUB-OLDT 4.0 was first developed at the Hub as a research tool. Almost immediately, there was a worldwide demand for its unique, transparent and anonymized data compilation capacity, which draws on real-life data but renders them as anonymous virtual avatars, so that the data are no longer relatable to any individual person. This was the key to processing individual data without infringement of personality rights or privacy. Far beyond the scientific community, venture capitalists, governments, and big corporations show an ardent interest in using the technology.

Open Simulation—what HUB-OLDT 4.0 can do

From the beginning, the goal was to provide access to the amazing potential of the simulator not only to a small circle of insiders. Following the examples of Open Access and Open Innovation of the early 2000s, its simulation experiments were intended to be not only available to the scientific community and international corporations but open to society at large: Open Simulation (os) was born. It was a logical next step in bringing science and society together again. Where, back then, Google, the NSA, Facebook, Amazon, and other corporations were dominating the market with individual-related personal data in ways unimaginable today, HUB-OLDT 4.0 offers a successful alternative. Where, back then, Google decided what we should think, and Amazon what we should buy, the humans of 2038 have a scientifically and socio-politically meaningful foundation available to them. Their decisions and actions are assigned a new space—that of simulation—in which democratic participation is practiced in an entirely different and freer way than just twenty years ago.

Der Simulator ermöglichte es, die Verwaltung eines Landes auf eine völlig neue Basis zu stellen. Sie erhält ein Navigationssystem in die nahe Zukunft, das die verschlungenen, meist nichtlinearen Wege zu den angepeilten Zielen anzeigt, die wiederum in ihrer Vernetzung und in ihren Widersprüchen in einer komplexen gesellschaftlichen Wirklichkeit gezeigt werden. Mit der einhergehenden Transparenz und der Möglichkeit, schlechte Lösungen und falsche Versprechen zu identifizieren, wurde der Siegeszug von »Fake News« und populistischer Tendenzen entscheidend eingebremst, die in den 2010er-Jahren zu einem erheblichen Problem für faktengetriebene Politik geworden waren. Demokratischen Entscheidungen wurde eine faktenbasierte Grundlage gegeben, die allgemein akzeptiert wurde. Das Vertrauen in demokratische Institutionen wurde wiederhergestellt, als die Menschen erkannten, dass die virtuellen Wirklichkeiten eines HUB-OLDT 4.0 komplexe Zusammenhänge sichtbar und ihre Konsequenzen verstehbar machten.

Das Versprechen einer authentischen Beteiligung der Zivilgesellschaft durch Open Simulation war kein leichtfertiges. Zu oft schon wurde ein technologisch fundiertes Empowerment angekündigt, zu oft eine aktive und informierte Beteiligung der Bevölkerung heraufbeschworen. Andererseits war es 2018 offenkundig, dass das Vertrauen in öffentliche Institutionen weiter schwand, während das Misstrauen gegenüber den Eliten wuchs. Es war und ist kein Allheilmittel für eine schwächelnde Demokratie. Wenn sie jedoch frühzeitig in eine zukunftsweisende Bildungspolitik eingebettet wird, kann spielerisch, d. h. durch Simulation, gelernt werden, dass die Diskussion von gesellschaftspolitisch relevanten Themen zum Nutzen der Demokratie auch auf einem höheren Niveau geführt werden kann.

Für Politik und Verwaltung kann die aktive Einbindung der Bevölkerung in OS – vom Gesundheitswesen zur Sicherheit, von der Bildung zur Gestaltung des urbanen Raums, von der Mobilität zur demografischen Entwicklung – zusätzliche Daten in sinnvoller und nutzbarer Form generieren. Menschen liefern dann ihre Daten, die durch das Fitnessband oder über ihre Kreditkarten gewonnen werden, nicht nur an Google, die NSA und Amazon, sondern werden mündige Vertragspartner einer Verwaltung, die ihnen als Gegenleistung ihre aktive Einbindung in das Simulationsexperiment zusichert.

Innovation wird anders organisiert

Innovation passt sich den gesellschaftlichen Anforderungen und Möglichkeiten laufend an. Sie ist längst zu einem komplexen Prozess geworden. Kreative Ideen und Menschen, die sie erzeugen, werden auch in Zukunft unentbehrlich sein. Der Kampf um intelligente Köpfe nimmt weiter zu. Doch gibt es weit mehr Ideen, als sich je in der realen Welt umsetzen lassen. HUB-OLDT 4.0 gelingt es heute, die aus unterschiedlichen Quellen stammenden Ideen nach bestimmten Kriterien zu bewerten und zu selektieren. Die Ideen lassen sich dann in dezentralisierte Innovationsnetzwerke einspeisen, deren Partner in engem Austausch miteinander stehen. Die Umsetzung von Ideen erfordert nach wie vor ein beträchtliches technisches Know-how sowie ein Verständnis der Märkte und der bestehenden oder beabsichtigten Regularien. In all diese Schritte sind die am globalen Innovationsnetzwerk »Industrie 19.0« beteiligten Firmen in Echtzeit eingebunden. Sie können unterschiedliche Strategien

The simulator made it possible to put the administration of a country on an entirely new basis with the help of twenty-first-century technologies, equipping it with a navigation system into the near future. This system provides direction along the tangled, mostly nonlinear paths towards the targeted goals, which in turn are outlined in their interrelatedness and confusions in a complex social reality. The entailed transparency as well as the possibility of identifying bad solutions and false promises decisively helped push back the triumphant advances of “fake news” and populist tendencies, which had posed a substantial problem standing in the way of fact-driven policymaking in the 2010s. Democratic decisions were put on a fact-based foundation that was generally accepted. Trust in democratic institutions was restored when people realized that the virtual realities of HUB-OLDT 4.0 made complex connections visible and their consequences comprehensible.

The promise of the civil society’s authentic participation through Open Simulation was not frivolously made. Too many times, technology-based empowerment had been heralded; too many times, active and informed citizen participation had been emptily invoked. On the other hand, it had become all too obvious in 2018 that trust in public institutions was on the wane and distrust against the elites was rising. OS was not, and is not, a nostrum for a faltering democracy. However, if embedded in future-oriented educational policies early on, it enables learning playfully—i. e., through simulation—that the discussion of relevant sociopolitical issues can also be taken to a higher level for the benefit of democracy.

For policymaking and administration, active involvement of the population in OS—from the health system to security, from education to urban planning and design, from mobility to demographic development—can generate additional data in a meaningful and usable form. People will then supply their data collected through their fitness trackers and credit cards not only to Google, the NSA, and Amazon, but will become voluntary contract partners of an administration that, in return, guarantees them active involvement in the global simulation experiment.

Reorganizing innovation

Innovation continuously adapts to societal requirements and possibilities. It has long become a complex process. Creative ideas and people to produce them will be indispensable also in the future. The battle for intelligent minds is increasing. Yet, there are far more ideas than can ever be implemented in the real world. HUB-OLDT 4.0 is able today to evaluate and select ideas from different sources according to specific criteria. Those ideas can then be entered into decentralized innovation networks whose partners are in close interchange with one another. The implementation of ideas still requires considerable technical expertise as well as an understanding of markets and of existing and intended regulatory mechanisms. The companies participating in the global “Industry 19.0” innovation network are involved in all these steps in real time. They can try out, experimentally expand or restrict differ-

ausprobieren, experimentell erweitern oder einschränken. Um den Firmen die Möglichkeit zu geben, ihre intellektuellen Eigentumsrechte zu schützen, folgt HUB-OLDT 4.0 weiterhin einem zweistufigen Verfahren, das schon vor Jahrzehnten am MIT Media Lab praktiziert wurde. Der erste Schritt ist völlig offen. Erst im zweiten erwirbt eine Firma das Recht, den Weg zum Patentschutz zu gehen.

Mit HUB-OLDT 4.0 tritt die Simulation aus dem Schatten geheim gehaltener Daten heraus. Jeder Schritt, den die Maschine und ihre lernenden Algorithmen machen, wird offen nachvollziehbar. Der leitende Grundsatz von os, so utopisch er nach wie vor scheinen mag, muss jedoch vor Missbrauch geschützt werden. Daher müssen sich alle Nutzer und Käufer von HUB-OLDT 4.0 vertraglich verpflichten, die Prinzipien von os einzuhalten. Darunter fällt das Eigentumsprinzip persönlicher Daten, das besagt, dass Individuen über uneingeschränkte Kontrolle über ihre Daten verfügen, was heute mit Blockchain-Technologien technisch längst zuverlässig umgesetzt ist.

Die schöpferische Kraft der Veränderung

HUB-OLDT 4.0 ist keine der vielgepriesenen oder gefürchteten disruptiven Technologien. Der Simulator ist auch kein den letzten Stand der Digitalisierung nutzendes Allheilmittel, um alle Probleme zu lösen. Er steht vielmehr für den Versuch, den Menschen die schöpferische Kraft der Veränderung – und nicht nur jene der Zerstörung – zugänglich zu machen. An konkreten Beispielen aus der praktischen Lebenswelt der Menschen lässt sich zeigen, dass es möglich ist, deren Vorstellungskraft und planbaren Handlungsspielraum durch Simulation dramatisch zu erweitern. Menschen werden ermutigt, von passiven Datenlieferanten für Großkonzerne, die ihnen anschließend Produkte und Dienstleistungen verkaufen, zu interaktiven und mündigen Partnern zu werden. Die schöpferische Kraft der Veränderung liegt im Prozess der Aneignung, die zur Gestaltung eines Möglichkeitsraums werden soll. Der Weg, der dorthin führt, geht über die Frage des Eigentums an Daten hinaus. Um den Möglichkeitsraum gestalten zu können, müssen die Kenntnisse, Erfahrungen und Interpretationen der Menschen ebenso einfließen wie die quantitative Erhebung eines breiten Verhaltensspektrums.

Künstliche Intelligenz bedeutet auch 2038 nicht, das eigene Denken auszuschalten oder einem Algorithmus zu überantworten. Zu entscheiden, wie und wofür eine Maschine einzusetzen ist und wie und wann das, was sie liefert, zu interpretieren und weiter zu nutzen ist, setzt unabhängiges Denken, setzt menschliche und wissenschaftliche Intelligenz voraus. Das wird uns HUB-OLDT 4.0 auch weiterhin nicht abnehmen. ✖

Der Text ist eine gekürzte Version des Kapitels »Ein Simulator verändert die Welt«, erschienen in: Hannes Androsch und Peter Pelinka (Hg.), *Zukunft. Erkennen | Gestalten. Perspektiven einer neuen Welt*, Wien 2018, S. 269–278.

ent strategies. In order to enable companies to protect their intellectual property rights, HUB-OLDT 4.0 continues to follow a two-step procedure already practiced decades ago at MIT. The first step is completely open. It is only in the second step that the company acquires the right to obtain patent protection.

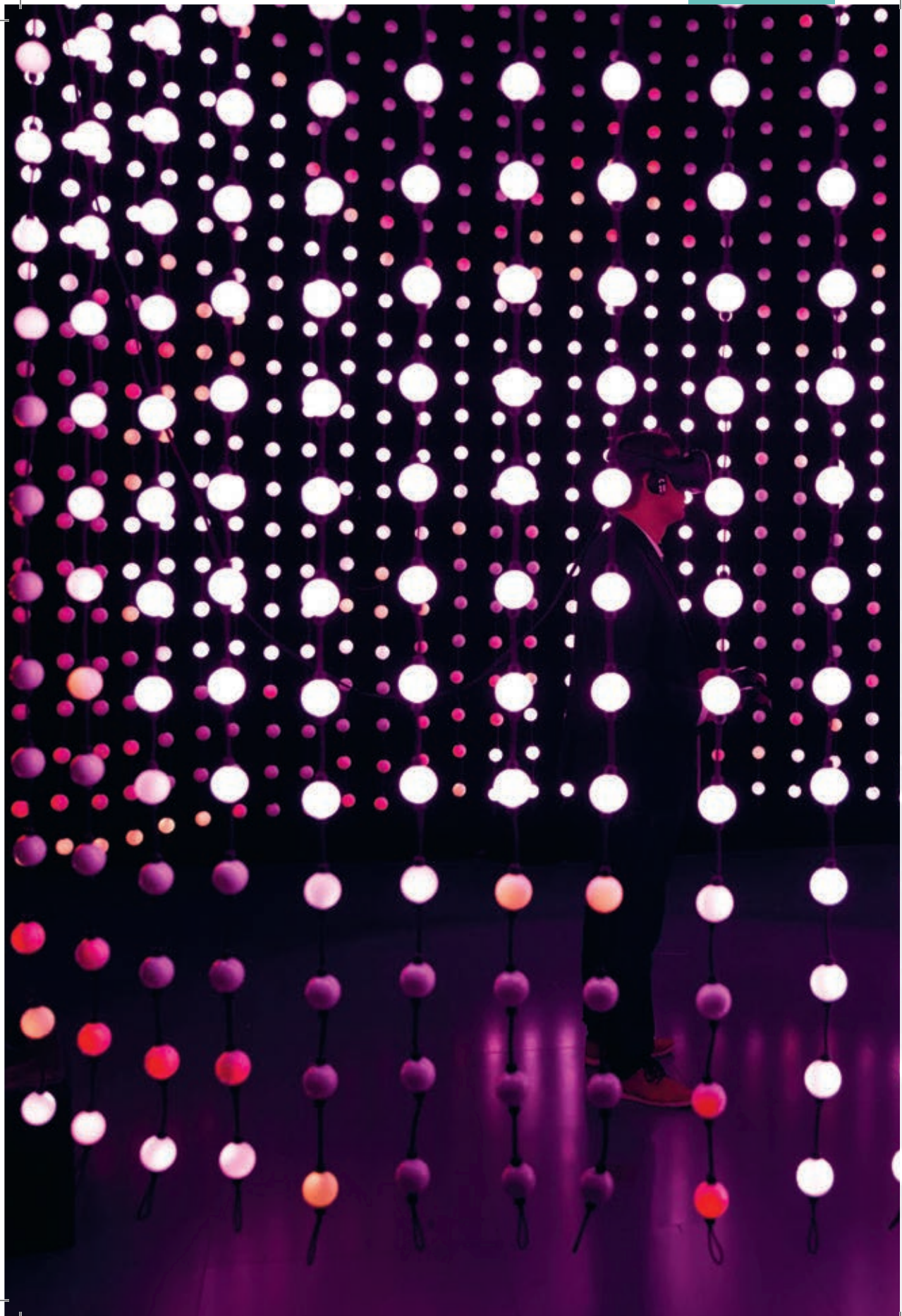
With HUB-OLDT 4.0, simulation steps out of the shadows of data secrecy. Each step taken by the machine and its learning algorithms is openly traceable. However, the guiding principle of os, utopian as it may still seem, has to be protected against abuse. Therefore, all users and buyers of HUB-OLDT 4.0 must contractually commit themselves to observe and comply with the principles of os. This includes the principle of the proprietary nature of personal data, meaning that individuals shall have full control over their data, something that today has long found reliable technical implementation using blockchain technologies.

The creative force change

HUB-OLDT 4.0 is not one of the much praised or feared disruptive technologies. Nor is the simulator a Swiss-knife universal problem solver using cutting-edge digital technology. Rather, it is an attempt to make the creative force of change—and not only of destruction—accessible to people. Using concrete examples from people’s practical everyday lives, it can be demonstrated that it is possible to dramatically expand their imagination and plannable scope of action through simulation. People are encouraged to develop from passive suppliers of data to big corporations, which then sell goods and services to them, into interactive and responsible partners. The creative force of change lies in the process of appropriation, which is supposed to turn into the unfolding of a space of possibility. The path that leads there extends beyond questions of data ownership. To unfold this space of possibility, there must be an input of individual human knowledge, experience, and interpretations as well as a quantitative survey of a broad range of behaviors.

Even in 2038, artificial intelligence does not mean shutting down one’s own thinking or leaving it all to an algorithm. Deciding how and what for to use a machine, and how and when to interpret and utilize what it supplies, still requires independent thought, human and scientific intelligence. That is something that HUB-OLDT 4.0 will not be able to take over from us. ✕

The text is an abridged version of a chapter entitled “A Simulator Changes the World,” published in Hannes Androsch and Peter Pelinka, eds., *Zukunft. Erkennen | Gestalten. Perspektiven einer neuen Welt* (Vienna, 2018), 269–278.



T
K

Ir
ir
ke
V
w
S
so
e
ü
ke
G
si
p
w
B
e
e
E
n
m
d
si
E
o
B

Ir
ci
n
ti
“
so
a
p
s
tä
s
v
e
u
o
o
ie
b
W
sc
at
ti

—
©
F

Training 2038 Kitchen Budapest (HU)

Im Gegensatz zu der in vielen Rechenprotokollen zu beobachtenden Verschleierung entwirft »Training 2038« ein Szenario, in dem Menschen eine Stimme erhalten und Feedback über Systeme liefern können, die gerade erst Gestalt annehmen. Im sicheren Raum einer privaten VR-Erfahrung wird eine umfangreiche Befragung in Form eines Dialogs zwischen einem sogenannten Embodied Conversational Chatbot und einem menschlichen User durchgeführt. Wie würde sich ein automatisierter Entscheidungsbaum ohne moralisches Bewusstsein verhalten?

In contrast to the perceived obfuscation in numerous computational protocols today, "Training 2038" stages a scenario where people are given a voice to provide feedback on systems that are just taking shape. In the safe space of a private VR experience, an extensive survey is underway in the form of a dialog played out between an embodied conversational bot and a human user. With no moral conscience, how would an automated decision tree behave?

© Ars Electronica/
Florian Voggeneder



Hermann Hauser im Gespräch mit Martin Kugler

»Wir brauchen Zugang zu den Daten, sonst können wir keine Forschung betreiben«



In den Bereichen Spracherkennung und autonome Fahrzeuge ist künstliche Intelligenz bereits drauf und dran, ganze Industriezweige umzukrempeln, stellt Hermann Hauser fest. Nachdem dank Deep Learning die Klassifizierungs- und Regressionsprobleme im Wesentlichen gelöst sind, sieht er die Automatisierung von Entscheidungen als nächsten großen Schritt.

Welche Rolle spielt künstliche Intelligenz (KI) schon heute in der Wirtschaft, und wie sieht aus Ihrer Perspektive die Vision für die nächsten Jahre aus?

Hermann Hauser: Die Menschen vergessen oft, dass KI ja schon seit Langem unter uns ist. Wir verwenden sie bereits täglich, zum Beispiel bei der Spracherkennung, die wirklich schon ausgezeichnet ist – teilweise schon wesentlich besser als die des Menschen. Auch die Objekterkennung ist bereits besser als die des Menschen. Das hängt damit zusammen, dass große Datensätze verfügbar sind, auf denen maschinelles Lernen aufbauen kann. Es gibt dazu ein schönes Beispiel vom Beginn der Spracherkennung beim iPhone: Burschen aus Glasgow haben sich bei Apple beschwert, dass das iPhone den Glasgower Akzent nicht versteht. Das ist ja auch nachvollziehbar, weil dort ein sehr ausgeprägter Dialekt gesprochen wird. Die Antwort von Apple lautete: Wartet ein paar Wochen! Der Grund, warum sie das mit dieser Bestimmtheit sagen konnten, war, dass jedes Mal, wenn ein Mensch aus Glasgow Siri verwendet hat, die Sprachdaten an Apple geliefert wurden. So konnte Apple eine Datenbasis erstellen, und die Spracherkennung funktioniert jetzt auch in Glasgow. Das gilt für alle Sprachen der Welt.

Was bedeutet das für die Wirtschaft?

HH: Da die Spracherkennung nun so gut ist, gibt es viele Möglichkeiten, Aufgaben über Spracherkennung zu erledigen. Überdies sind die modernen Speaker, die viele Menschen schon zu Hause stehen haben, neue Produktparten, die es

Hermann Hauser, Mitglied der Royal Society und des Institute of Physics der Royal Academy of Engineering sowie Ehrenmitglied des King's College, ist ein weltweit erfolgreicher Computer- und Venture-Capital-Unternehmer. Seine Firma Acorn entwickelte ab 1978 den ARM-Prozessor. 1997 gründete Hermann Hauser Amadeus Capital Partners und später das Hauser Forum in Cambridge, in dem inzwischen neben anderen Firmen ideaSpace, Cambridge Enterprise und Cambridge Network Quartier bezogen haben. 2016 gründete er das I.E.C.T. – Institute for Entrepreneurship Cambridge – Tirol, um die Start-up-Landschaft im alpinen Raum zu stärken.

Hermann Hauser, Fellow of the Royal Society and the Institute of Physics of the Royal Academy of Engineering as well as Honorary Fellow of King's College, is a successful computer and venture capital entrepreneur. His company Acorn developed the ARM processor from 1978 on. Hermann Hauser founded Amadeus Capital Partners in 1997 and later set up the Hauser Forum in Cambridge, which accommodates ideaSpace, Cambridge Enterprise, Cambridge Network, and other firms today. In 2016, he founded I.E.C.T. - Institute for Entrepreneurship Cambridge - Tirol to support the start-up scene in the Alpine region.

Hermann Hauser in an interview with Martin Kugler

“We need access to data, otherwise we cannot do research”

In the fields of speech recognition and autonomous vehicles, artificial intelligence is on the verge of turning entire industries inside out, Hermann Hauser points out. As problems of classification and regression have essentially been solved thanks to deep learning, he believes that the next major step will be the automation of decisions.

What role does artificial intelligence (AI) already play in industries and what's the vision for the years to come from your perspective?

Hermann Hauser: People frequently tend to forget that AI has been part of our lives for a long time. For example, it's already used on a daily basis when it comes to speech recognition, which has improved enormously—and is partly even better than that of humans. The same holds true for object recognition. This is because machine learning can build on huge datasets. There is a wonderful example from the beginnings of speech recognition concerning the iPhone: Fellows from Glasgow complained to Apple that the iPhone didn't recognize their accent—which is entirely understandable, as their dialect is indeed very special. Apple's response: Wait a couple of weeks! The reason for this confidence was that each time someone from Glasgow was using Siri, the speech data was sent to Apple. This allowed Apple to compile a database, with the result that speech recognition now also works in Glasgow. This goes for all of the world's languages.

How will this impact on the economy?

HH: As speech recognition has improved so much, there are now numerous possibilities to carry out tasks via speech recognition. Moreover, the modern speakers that can meanwhile be found in the homes of many people represent a new product line that did not exist earlier. One of the crucial turning points will be autonomous driving, which is going to turn the entire automobile industry inside out. Many other industries will follow in its wake.

bisher nicht gab. Einer der größten Einschnitte wird aber das autonome Fahren sein, das die ganze Autoindustrie umkrempeln wird. Das wird auch in vielen anderen Branchen so sein.

Warum sind Sie da so sicher?

HH: Es gab sechs Computerwellen, die letzte war die PC-Welle. Es war bei allen sechs Wellen bisher so, dass jene Firmen, die eine Welle dominierten, immer die nächste Welle verschlafen haben. Es konnte sich zum Beispiel niemand vorstellen, dass Microsoft und Intel, die die PC-Welle dominiert hatten, bei Smartphones keine Rolle spielen würden. Ich habe lange Zeit geglaubt, dass das vielleicht nur auf die Computerindustrie zutrifft. Mir ist aber inzwischen klar geworden, dass das für alle Industriesparten zutrifft – und zwar unter den folgenden zwei Umständen: dass es erstens eine disruptive neue Technologie gibt, die zweitens mit einem neuen Geschäftsmodell verbunden ist. Der Grund dafür, dass meine Firma ARM (Advanced RISC Machines) Intel vollkommen aus dem Mikroprozessorbereich verdrängt hat, war die Kombination eines technischen Durchbruchs – in unserem Fall RISC und Low Power – und der Umstellung des Businessmodells. Wir lizenzieren die Mikroprozessoren und bauen sie nicht selber. Wir haben 500 Lizenznehmer und einen 95-prozentigen Marktanteil bei Smartphones. Wir verkaufen zur Zeit 20-mal mehr Mikroprozessoren als Intel. Der Grund, warum Intel das verschlafen hat, ist, dass sich Intel gegen ARM nicht wehren konnte, weil sein Produkt ein Chip war. Smartphone-Produzenten kaufen keine Chips, sondern wollen Lizenzen haben. Intel konnte das Geschäftsmodell nicht ändern. Die Lizenzgebühren machen nur ein paar Prozent eines Chips aus: Daher hätte es keinen Sinn für Lizenznehmer, eine eigene Architektur zu bauen. Die Lizenzgebühr ist im Vergleich zu den Investitionen in eine eigene Architektur sehr gering. Außerdem haben Lizenznehmer kein Entwicklungsrisiko.

Gibt es neben der Spracherkennung und dem autonomen Fahren weitere Bereiche, wo KI in absehbarer Zeit eine große Rolle spielen wird?

HH: Ja, zum Beispiel bei Recommender-Systemen, die schon jetzt ziemlich großen Einfluss auf den Umsatz von Onlinehändlern haben. Der wichtigste Punkt bei KI ist aber, dass das Werkzeug so generell ist: Es ist das leistungskräftigste Werkzeug, das wir als Menschen jemals erfunden haben. KI ist vielleicht vergleichbar mit so grundlegenden Innovationen wie der Elektrizität und wird in allen Wirtschaftsbereichen wichtig werden. Das autonome Fahren bringt nun eine unmittelbare Umstrukturierung eines ganz großen Industriezweiges mit sich. KI ist aber nicht die einzige neue

How can you be so sure?

HH: There have been six digital waves so far, the last one having been the PC wave. For all of the six previous waves it can be said that those companies dominating one wave missed out on the next. For instance, no one would have imagined that Microsoft and Intel, having dominated the PC wave, wouldn't be playing a role in the smartphone business. For a long time I assumed that this might only apply to the computer industry. In the meantime, though, I've realized that this holds true for all the industries—namely under two circumstances: first, that there is a disruptive new technology, and second, that it is linked to a new business model. The reason why my company ARM (Advanced RISC Machines) succeeded in squeezing Intel out of the microprocessor market completely was the combination of a technological breakthrough—in our case RISC and low power—and a revision of the business model. We license the production of microprocessors instead of building them ourselves. We have 500 licensees and a 95-percent market share in the smartphone business. We are currently selling twenty times as many microprocessors as Intel. The reason why Intel slept through this development was that they couldn't compete with ARM because their product was a chip. Smartphone producers don't buy chips but prefer to hold licenses. Intel failed to modify its business model. License fees only account for a few percent of what a chip costs. It would therefore be superfluous for licensees to build their own architecture. The license fee is very modest compared to investing in your own architecture. In addition, licensees don't take any risks when it comes to development.

Are there any other fields besides speech recognition and autonomous driving where AI will be playing a major role in the near future?

HH: Yes, recommender systems, for instance, which already significantly affect the sales of online retailers. The most important point, however, is that AI is such a universal tool: it's the most efficient tool we humans have ever invented. AI can probably be compared with such fundamental innovations as electricity and will be important in all industrial sectors. Autonomous driving is now bringing about an instant reorganization of an extremely powerful large industry. But AI is not the only innovative technology that will cause such profound changes. In my opinion, there are three such technologies: first, AI; second, the blockchain technology with its smart contracts, which will ensure the auto-

Technologie, die so große Umwälzungen bewirken wird. Meiner Meinung nach gibt es drei solche Technologien: erstens KI, zweitens die Blockchain-Technologie mit den Smart Contracts, die viele Geschäftsprozesse automatisieren werden, und drittens die synthetische Biologie, die zu einer Revolütierung der Life Sciences führen wird. Im Finanzwesen wird die Blockchain genauso fundamental sein wie die KI bei Autos.

In welchem Stadium des Innovationszyklus befindet sich die KI zur Zeit? Ist sie noch im Pionierstadium? Haben wir das »Tal der Tränen« schon durchschritten? Oder wird die Technologie schon breit ausgerollt?

HH: Die Technologie wird noch nicht breit ausgerollt. Ich glaube, dass wir gerade das »Tal der Tränen« hinter uns haben. Ich hätte in den späten 1970er-Jahren meine Dissertation fast in KI und nicht in Physik gemacht, weil ich unbedingt das World-Chess-Champion-Programm mitschreiben wollte. Das sollte damals im Herbst des Jahres fertiggestellt werden. Es hat dann noch einmal 20 Jahre gedauert. Aber die Grundideen waren schon damals vorhanden, und die unterscheiden sich gar nicht so sehr von den Ideen heute. Geoffrey Hinton, einer der Begründer der modernen KI, hat drei Gründe genannt, warum das jetzt endlich funktioniert – nach diesen vielen Jahren des KI-Winters. Die drei Gründe nach zunehmender Wichtigkeit: Wir haben mehr Computerleistung. Die Algorithmen sind etwas besser geworden – die Grundideen gab es wie gesagt schon vor 30 Jahren. Am wichtigsten ist aber, dass es jetzt große Datensätze gibt, mit denen maschinelles Lernen realisierbar geworden ist.

Wo geschieht die Weiterentwicklung der KI? Sind Start-up-Unternehmen die treibende Kraft? Oder ist es die Industrie? Oder die akademische Welt?

HH: Es gibt immer noch akademische Durchbrüche. Ein Beispiel sind die Gaussian Processes in Cambridge, wo wir darauf aufbauend eine Firma, die PROWLER.io, gegründet haben. Das ist ein neuer Ansatz, der meiner Meinung nach der nächste Schritt in der KI ist – nach Deep Learning. Mit Deep Learning haben wir die Klassifizierungs- und Regressionsprobleme im Wesentlichen gelöst. Wir haben zum Beispiel mit Swarovski ein Projekt gemacht, bei dem in das Bild im Fernglas eingeblendet wird, welchen Vogel man gerade sieht. Das funktioniert wunderbar. Die Klassifizierung ist aber nicht die zentrale Frage. Meine Lieblingsdefinition von Intelligenz ist: »Knowing what to do next.« Es geht um Entscheidungsfindung. Unser Leben ist eine Kette von Entscheidungen, man muss jeden Augenblick und auf allen Ebenen Entscheidungen treffen. Welchen Laut produzieren wir in der nächsten Millisekunde? Was ist das nächste Wort? Der nächste Satz? Und

mation of countless business transactions; and third, synthetic biology, which will lead to a revolution in the life sciences. In finance, the blockchain will be just as fundamental as AI in the automobile sector.

What's the current stage of innovation of AI? Is it still in its pioneering stage? Have we already passed through the "trough of disillusionment"? Or has the technology already been launched on a large scale?

HH: No, not on a large scale. I think we've left the "trough of disillusionment" behind us. In the late 1970s, I was close to making AI the subject of my dissertation—in-
stead of physics—because I badly wanted to participate in writing the World Chess Champion program, which should have been completed by fall that year. Eventually it took another twenty years. But the basic ideas already existed then and were not so different from today's. Geoffrey Hinton, one of the originators of modern AI, gave three reasons why it finally functioned—after those many years of AI winter: our computers have become more efficient; algorithms have improved—as I said, the basic ideas already existed thirty years ago; but most importantly, we now have huge datasets at our disposal through which machine learning has become feasible.

In what spheres is the development of AI being pushed ahead? Are startup companies the driving force? Or industry? Or academia?

HH: There are still academic breakthroughs, such as research into the Gaussian processes undertaken in Cambridge, which also prompted the foundation of the company PROWLER.io. This is a new approach, which I believe will be the next step in AI, coming after deep learning. Through deep learning we have essentially solved the problems of classification and regression. For instance, we did a project with Swarovski on binoculars informing you on a display about what species of bird you are watching at the moment. This works just wonderfully. Yet classification is not the central problem. My favorite definition of intelligence is: "knowing what to do next." It's all about decision-making. Life is a chain of decisions; you have to make decisions every moment and on all kinds of levels. What sound are you going to produce in the next millisecond? What will be your next word? Your next sentence? What are we going to do after our talk? It seems to me that tracking such processes with computers will be the next step of AI. Decisions play a central role in many different fields, such as when it comes to so-called QUBO problems: When a cab has dropped you off, its driver must decide where to go

was machen wir nach unserem Gespräch? Das mit dem Computer nachzuvollziehen, scheint mir der nächste Schritt in der KI zu sein. Entscheidungen spielen auf sehr vielen verschiedenen Gebieten eine zentrale Rolle. Zum Beispiel bei einem sogenannten QUBO-Problem: Wenn eine Taxifahrt abgeschlossen ist, muss man sich entscheiden, wohin das leere Taxi fahren soll, damit es die beste Chance hat, den nächsten Kunden zu bekommen. Das ist ein ziemlich kompliziertes Problem, zu dem man drei Komponenten braucht: *probabilistic modelling*, *reinforcement learning* und *multi-agent decision-making*. PROWLER.io ist die erste Firma, die alle drei Komponenten zusammengebracht hat.

Trifft das klassische Innovationsmodell – dass eine Idee aus der akademischen Welt kommt, dann ein Start-up gegründet wird und die Innovation schließlich von der Industrie breit ausgerollt wird – auch auf KI zu?

HH: Nein, denn bei GAFA (Google, Apple, Facebook und Amazon) gibt es ein neues Phänomen: Die Unternehmen haben so viel Geld, dass sie praktisch wie Universitäten funktionieren können.

Das erinnert an Zeiten wie vor 70 Jahren, als es noch die riesigen Bell Laboratories gab...

HH: Ja, das gab es schon einmal. Da sind seinerzeit alle hingepilgert. Es war wirklich schade, dass sie zugesperrt wurden.

Das kommt aber nun wieder?

HH: Ja, Google's DeepMind in London zum Beispiel ist eines der aktivsten KI-Forschungslabors der Welt. Dort arbeiten, glaube ich, 600 Wissenschaftler.

Wer investiert vor allem in KI?

HH: Im Augenblick alle. Ich bekomme keinen Businessplan, der nicht eine KI-Sparte hat. Darunter gibt es einige wirklich tolle Gruppen.

Wie groß ist in diesem Bereich eigentlich die Ausfallsrate? Ist sie ähnlich hoch wie etwa im Biotechbereich?

HH: Das kommt darauf an, in welchem Stadium ein Start-up ist. In der Anlaufphase liegt die Ausfallsrate bei über 50 Prozent. In der »Series A«-Runde ist es schon wesentlich besser: Die Ausfallsquote liegt dort vielleicht bei 30 Prozent. Bei erfahrenen Gründern ist die Chance noch wesentlich höher. Auf jeden Fall ist die Ausfallsquote niedriger als im Biotechbereich. Unterzieht man die Technologie einer guten Due-Diligence-Prüfung, kommt es nur selten zu Enttäuschungen. Im KI-Bereich und allgemein in den

next to have the best chance to find a new passenger. This is a rather complex problem, which involves three components: *probabilistic modeling*, *reinforcement learning*, and *multi-agent decision-making*. PROWLER.io is the first company to have brought together all of these three components.

Does the conventional innovation model—an idea hatched up by academia, the establishment of a startup, and industry’s launch of the resulting innovation on a broad scale—also apply to AI?

HH: No, because in the case of GAFA (Google, Apple, Facebook, and Amazon) a new phenomenon can be observed: these companies have so much money that they are in a position to virtually operate like universities.

This reminds me of the times seventy years ago, when the huge Bell Laboratories still existed...

HH: Yes, we had that once before. In those days, everyone pilgrimaged there. It’s really a pity that they were closed down.

But this is coming again now?

HH: Yes, Google’s DeepMind in London, for example, is one of the busiest AI research laboratories worldwide, employing, I think, some 600 scientists.

Who are currently the major investors in AI?

HH: At the moment, everyone is investing in AI. I don’t see a single business plan that doesn’t take AI into account. And there are some really amazing groups out there.

What about the failure rate in this industry? Is it similarly high as in biotechnology?

HH: This depends on the phase of development a startup is in. In the seed stage, the failure rate is more than fifty percent. The series A round looks much better already: the failure rate seems to be around thirty percent. When experienced founders are involved, chances are considerably higher. In any case, the failure rate is lower than in biotechnology. If the technology has undergone a thorough due-diligence examination, disappointment is rare. In AI and in information and communication technology (ICT) in general, the technology will work most of the time if you have a good team. It probably does not work as well as the people think, so that it may take one or two years longer. But normally you can be confi-

Informations- und Kommunikationstechnologien (IKT) ist es meist so, dass die Technologie funktioniert, wenn man ein gutes Team hat. Vielleicht funktioniert sie nicht ganz so gut, wie die Leute denken, und vielleicht dauert es ein oder zwei Jahre länger. Aber im Normalfall kann man sich darauf verlassen, dass die Technologie funktioniert. Im Biotechbereich ist es umgekehrt: Dort ist die Ausfallsrate in den späten klinischen Phasen relativ hoch. Ich sehe die Zukunft des Gesundheitsbereichs aber nicht in der Entwicklung von Medikamenten, sondern glaube an eine große Umstellung des Geschäftsmodells: Man wird sich nicht mehr auf Medikamente für kranke Leute konzentrieren, sondern auf Menschen, die gesund sind und gesund erhalten werden sollen.

Dafür sind IKT und KI wesentliche Bausteine, oder nicht?

HH: Ja. Das ist die wichtigste Verwendung von Datensätzen über den menschlichen Körper, die man jetzt durch Biotech und durch Sensoren aufbaut. Diese Sätze sollen dazu verwendet werden, Prognosen zu stellen – und zwar schon lange bevor jemand eine Krankheit entwickelt. Diese Systeme sollen uns sagen: Pass auf, du hast da eine Prädestination für diese bestimmte Krankheit! Wir schlagen dir folgende Umstellung deines Lebens oder deiner Ernährung vor.

Gibt es dafür schon Geschäftsmodelle?

HH: Im Augenblick sind die Sensoren und die Produktion von Datensätzen durch Genomik, Proteomik und alle anderen -omics das Geschäftsmodell. Man braucht wahrscheinlich nur ein oder zwei Dutzend Grundparameter des Körpers alle paar Sekunden rund um die Uhr zu messen – etwa Blutdruck, Puls, Atemfrequenz, Blutgase und so weiter –, dann bekommt man einen Gesundheitsvektor, der für jeden Menschen eine Art Baseline darstellt. Wenn die aktuellen Messwerte zu weit davon abweichen, kann man ein Feedback vornehmen. Zum Beispiel: Das war jetzt genug Schokolade. Oder Alkohol.

Was verspricht sich jemand davon, wenn er in KI investiert?

HH: In erster Linie natürlich Profit.

Aber wollen die Menschen nicht auch bei etwas Neuem dabei sein und die Welt der Zukunft mitgestalten? Will nicht auch die bestehende Industrie an den Innovationen teilhaben?

HH: Na eben nicht! Meine Prognose ist, dass es von den zehn größten Autofirmen der Welt in zehn Jahren drei bis fünf nicht mehr geben wird. Der Grund dafür ist ein doppelter: Einmal führen die Auswirkungen des autonomen Fahrens dazu, dass es eine radikale Reduktion der Anzahl der

dent that the technology will work. In biotechnology, it's the other way round: the failure rate is relatively high in the late clinical stages. However, I don't see the future of the health sector in the development of new drugs but believe in a profound reform of its business model: instead of concentrating on drugs for sick people, the focus must be on healthy people and on keeping them in good health.

ICT and AI are essential building blocks for this, aren't they?

HH: Yes. This will be the most important application of the human body datasets that are now being compiled with the aid of biotechnology and sensors. These sets will be used to make prognoses—that is, long before a person develops an illness. These systems are supposed to tell us: Watch out, you're predestined to develop this specific disease! We suggest the following changes in your lifestyle and diet.

Are there already business models for this application?

HH: At the moment, it's the sensors and the production of datasets through genomics, proteomics, and all the other -omics that make up the business model. We'll probably only need to measure one or two dozens of basic physical parameters—such as blood pressure, heart rate, respiratory rate, blood gases, etc.—every few seconds and around the clock. This will provide us with sufficient information to define a health vector that will represent a kind of baseline for each individual. If the current measurement data deviates from it excessively, feedback can be given. For example: You've had enough chocolate. Or alcohol.

What do people investing in AI expect for themselves?

HH: First of all, profit, of course.

But don't people feel the desire to be part of something new and contribute to the design of the future world? Don't the established industries too wish to participate in these innovations?

HH: That's exactly what they don't want! My prognosis is that in ten years' time three to five of the world's largest automobile manufacturers will no longer exist. There's a dual reason for this: First of all, autonomous driving will lead to a radical decrease in the number of cars. It's predicted that in some cities 80 percent fewer cars will be needed. The main problem will be a change in the sector's business system: there'll be a profound shift in dominance from the producers to the service

Autos geben wird, die wir brauchen. In manchen Städten, so die Prognose, werden wir 80 Prozent weniger Autos brauchen. Das Hauptproblem ist aber die Änderung des Geschäftssystems: Die Hauptmacht in dieser Sparte geht weg von den Produzenten hin zur Serviceindustrie. Junge Leute, wie zum Beispiel mein Sohn, wollen keinen Führerschein mehr machen. Er lebt in London, ist mit der U-Bahn glücklich, und wenn er ein Mädchen ausführen will, nimmt er ein Uber-Taxi. Für uns war Autofahren ein großer Spaß, die jungen Menschen heute haben aber keine Freude am Fahren, sondern eher Angst, etwas falsch zu machen. Das ist eine kulturelle Umstellung: vom Selber-ein-Benzinauto-Fahren zu einem Elektroauto, das autonom fährt.

Was bedeutet das in der Praxis?

HH: Als Konsument kauft man nicht mehr ein Auto um 20.000 oder 30.000 Euro und macht sich vorher lange Gedanken über das Aussehen, die PS oder den Verbrauch. Das neue Kriterium ist jetzt eine 10-Euro-Entscheidung: Welches Taxisystem verwende ich? Die Kriterien sind dabei ganz andere: Ist das Auto gleich da? Wie viel kostet das Taxi? Ist es bequem genug? Ob das Auto fesch aussieht oder viele PS hat, ist irrelevant. Diese Umstellung des Geschäftssystems bedeutet, dass die Macht in der Autoindustrie von den Herstellern auf die Dienstleister übergeht. Diese wollen große Stückzahlen von Autos einkaufen und mithilfe ihrer Technologie ihre Services am Markt anbieten. Sie können sich dafür den billigsten Anbieter eines Autos aussuchen. Die Frage ist nun: Können sich die Autohersteller dagegen wehren – oder geht es ihnen wie Intel bei ARM? Die Antwort auf diese Frage ist nicht klar. Die Autohersteller kennen die Zusammenhänge natürlich genau, und wie Beispiele zeigen, können sie auch eigene Services anbieten. Das Gegenargument ist aber, dass ein ganz wichtiger Teil der Entscheidungsfindung darin besteht, wie man das billigste Elektroauto kaufen kann. Wenn ein Autohersteller solche Services selbst anbieten will, wird er wohl seine eigenen Fahrzeuge dafür verwenden müssen. Er wird sich nicht den billigsten Produzenten aussuchen können.

industry. For instance, young people like my son are no longer so keen on getting their driver's license. He lives in London and is happy with the Underground, and when he wants to go out with a girl he books an Uber taxi. For us, driving a car meant a lot of fun, but today's youngsters don't enjoy driving, they rather fear that they might do something wrong. This is a cultural change: from driving a gasoline engine car yourself to an autonomously driving electric car.

What does this mean in practice?

HH: A consumer will no longer buy a car for 20,000 or 30,000 euros and be wracking his or her brain about its appearance, engine power, or consumption prior to the purchase. The new criterion boils down to a 10-euro decision: Which cab system should I use? The criteria are entirely different: Is the car going to pick me up at once? How much does the cab cost? Is it sufficiently convenient? The car's looks and horsepower are totally irrelevant. This change in the business system will result in a shift in control within the automobile industry from manufacturers to service providers. These providers seek to buy cars in large numbers and offer their services on the market with the aid of their technologies. For this, they can choose the most inexpensive car supplier. The point is: Will car manufacturers be able to fight this scenario—or will they end up as Intel when confronted with ARM? The answer is not yet clear. Of course, car manufacturers are entirely familiar with the situation and, as examples have shown, in a position to offer their own services. The counter-argument is that an essential part of the decision-making process is how to find out about buying the most inexpensive electric car. If car manufacturers decide to offer such services themselves, they will have to use their own vehicles. They won't have the chance to choose the cheapest producer.

Further important aspects of AI are privacy and security. We witnessed the Facebook data scandal this year, with data having obviously been abused for political ends. How does this influence developments in the field of AI?

HH: It's important that this debate is taking place now and that all of us are becoming aware of the fact that the business models of Google or Facebook are unfair. It has never been explicitly stated that they provide us with a search function or a social media portal free of charge in exchange for our data. Most people didn't know that their data was so important. This has now become

Weitere wichtige Aspekte von KI sind Privacy und Security. Wir haben heuer den Facebook-Skandal um Daten erlebt, die offenbar für politische Zwecke missbraucht wurden. Welchen Einfluss hat das auf die Entwicklung im Bereich der KI?

HH: Es ist sehr wichtig, dass diese Diskussion jetzt stattfindet, dass sich jeder dessen bewusst wird, dass die Geschäftsmodelle von Google oder Facebook unfair sind. Es wurde nie ausgesprochen, dass sie uns kostenlos eine Suchfunktion oder ein Social-Media-Portal dafür zur Verfügung stellen, dass wir ihnen unsere Daten liefern. Die meisten Leute haben nicht gewusst, dass ihre Daten wichtig sind. Das ist nun explizit geworden. Die neuen Regeln, die Europa nun im Rahmen der General Data Protection Regulation erlässt, sind sehr wichtig: Die Menschen sollen sich bewusst entscheiden können, welche Daten sie für diese Dienste hergeben wollen und welche nicht. Man darf das Kind aber nicht mit dem Bade ausschütten. In Österreich hatten wir heuer auch die Diskussion, ob Registerdaten und andere öffentliche Daten für die Forschung zugänglich gemacht werden sollen. Wir brauchen Zugang zu den Daten, sonst können wir keine Forschung betreiben. Solange die Daten anonymisiert sind, ist das meiner Meinung nach zu befürworten. Die große Diskussion ist, wie wir die Daten teilen. Das ist natürlich nicht einfach. Der Bevölkerung muss differenziert klargemacht werden, welche der vielen Daten für das Gemeinwohl wichtig sind. Nehmen wir zum Beispiel Daten über die Verkehrsdichte in der Stadt: Da ist jeder einzelne Autofahrer ein Datenpunkt. Um zu wissen, wie viele Autos auf einer Strecke unterwegs sind, ist es aber nicht wichtig, ob ein bestimmter Autofahrer zu einem bestimmten Zeitpunkt dort war, sondern nur, dass ein Auto dort war. Für die Verkehrsplanung ist nicht der Einzelne wichtig, sondern der Durchschnitt. Eine Lösung dafür ist die Anonymisierung.

Haben Sie, falls es zu sehr restriktiven Rahmenbedingungen beim Datenschutz kommen sollte, keine Sorge, dass das die Entwicklung im Bereich KI abwürgen könnte und wir auf einen neuen KI-Winter zusteuern?

HH: Nein, das glaube ich nicht, denn die Vorteile der Technologie und die Produktivitätssteigerungen durch sie sind sehr groß. Es wird bestimmt immer wieder holpern – wie jetzt bei Facebook –, aber diese Probleme werden sich lösen lassen.

clear. The new legislation Europe is going to pass within the General Data Protection Regulation is extremely important: people should be given the chance to make a conscious decision as to which data they will give in exchange for these services and which not. However, we should be careful not to throw out the baby with the bath water. Austria has seen discussions this year whether register data and other public data should be made available for research. We need access to data, otherwise we cannot do research. As long as the data has been anonymized, this should, in my opinion, be permitted. There's a heated debate going on now about how this data will be shared. It goes without saying that this is not easy. The population has to be informed in detail which of the huge amounts of data serve the common good. Take the data about a city's density of traffic, for example: every single driver is a valuable data point. To find out how many cars frequent a certain route, it's not necessary to know, however, if a particular driver was there at a specific point in time. What's important is to know that a car was there. For traffic planning, it's not the individual that counts, but an average value. The solution thus lies in depersonalization.

Don't you fear that an overly restrictive framework for data protection might throttle the development of AI and that we might be heading for another AI winter?

HH: No, I don't think so, because the technology's advantages and the increases in productivity are enormous. There'll always be obstacles and problems—such as with Facebook—but these can be solved.

It is said that the world is looking to Europe when it comes to the handling of data.

HH: Yes, Europe has done a really good job here and is the global leader with these regulations.

Es heißt, dass die Welt im Hinblick auf die Gestaltung des Rahmens im Umgang mit Daten auf Europa blickt.

HH: Ja, Europa hat da wirklich gute Arbeit geleistet und ist mit diesen Regelungen weltweit führend.

Weiters hört man immer wieder, dass die nächste Generation von IKT – von KI bis zum Quantencomputer – eine große Chance für Europa sei, seinen Stellenwert in der Welt wieder ausbauen zu können. Sehen sie das auch so?

HH: Ja, das sehe ich auch so. Daher arbeite ich auch beim European Innovation Council (EIC) mit; ich bin dort Chairman der High Level Group of Innovators. Im neuen Rahmenprogramm sind zwei Milliarden Euro pro Jahr für Innovationen in Europa vorgesehen, die genau diese Bereiche unterstützen sollen. In Europa haben wir immer noch ein großes Manko bei Risikokapital – es gibt ungefähr viermal weniger pro Kopf und Nase als in Amerika. Zudem haben wir Rahmenbedingungen, die es KMU wesentlich schwieriger machen zu reüssieren als in Amerika.

Wird China in diesem Spiel auch mitmischen?

HH: China ist im Augenblick die führende KI-Nation, noch vor Amerika. Es ist erstaunlich, wie schnell China das geschafft hat. Einer der Vorteile von China ist natürlich, dass die betreffenden Stellen wesentlich weniger Skrupel beim Umgang mit Daten haben.

Ist es in dieser Hinsicht nicht kontraproduktiv, wenn Europa beim Datenschutz wesentlich restriktiver wird?

HH: Europa muss wesentlich klarer werden. Man muss wirklich aufpassen, dass man das Kind nicht mit dem Bade ausschüttet. Denn die großen Datensätze sind eine Grundvoraussetzung für KI. Um KI wirklich gut arbeiten zu lassen, braucht man große anonymisierte Datensätze. Je besser ein System reguliert ist, desto höher ist die Qualität der Daten. ✕

It's also said that the next generation of ICT—from AI to quantum computing—will be a great chance for Europe to regain ground in the world. Do you agree?

HH: Yes, I do. This is why I also work for the European Innovation Council (EIC), where I'm chairman of the High Level Group of Innovators. The new framework program allots two billion euros per year to innovation in Europe to support exactly these areas. In Europe, we still suffer from a huge lack of venture capital—which is approximately four times less per capita than in the United States. Moreover, our framework conditions make it much more difficult for SMEs to be successful than in the United States.

Will China be part of the game?

HH: China is currently the leading AI nation, outperforming even the United States. It is amazing how fast China managed to achieve this. One of China's advantages is of course that the authorities in charge have considerable fewer qualms about using data.

Wouldn't it be counterproductive then if Europe became substantially more restrictive about data protection?

HH: Europe must take a clearer stance. As I said, we must really be careful not to throw out the baby with the bath water. Because these huge datasets are a fundamental prerequisite for AI. In order to be able to harness AI really efficiently, big anonymous datasets are needed. The better a system is regulated, the better the quality of the data will be. ✕

Gerald Reischl

Digital India: Mit Technologie zurück zur Weltmacht



Im Jahr 2060 werden in Indien nicht nur 1,7 Milliarden Menschen leben, sondern das Land soll dann auch (wieder) die größte Wirtschaftsmacht und Wissensgesellschaft der Welt sein. Mit einer noch nie dagewesenen Digitalisierungsoffensive will die Regierung die Infrastruktur und die Voraussetzungen schaffen, um dieses Ziel zu erreichen. Indien setzt dabei auf die Innovationskraft der eigenen erfolgreichen Unternehmen, den Veränderungs- und Gestaltungswillen der Bevölkerung und auf Zukunftsthemen wie Big Data, Blockchain, künstliche Intelligenz und das Internet der Dinge.

Indien will dorthin zurück, wo es bereits vor 300 Jahren war. Bis ins 18. Jahrhundert waren laut Analysen des 2010 verstorbenen britischen Ökonomen Angus Maddison Indien und China die größten Wirtschaftsmächte der Welt, ehe beide Nationen Anfang des 19. Jahrhunderts von Europa und dann, Mitte des 19. Jahrhunderts, auch von den Vereinigten Staaten überholt wurden. Seit der Jahrtausendwende geht es wieder stetig bergauf. Chinas immer wichtiger werdende Position im weltweiten Gefüge ist bekannt, Indien will es China jetzt nicht nur gleichtun, sondern den großen Nachbarn überflügeln.

Stärker als der Rivale China werden

Das Wirtschaftswachstum Indiens ist derzeit beachtlich, das Bruttoinlandsprodukt (BIP) ist 2017/18 um 6,75 Prozent gestiegen. Für das bis Ende März 2019 laufende Haushaltsjahr peilt die indische Regierung einen BIP-Zuwachs von 7,0 bis 7,5 Prozent an – übrigens um 0,5 bis 0,7 Prozentpunkte mehr als der Rivale China. Indien ist damit eine der am schnellsten wachsenden Wirtschaftsmächte. Verantwortlich dafür sind Initiativen wie »Digital India« oder »Make in India« (damit will Indien zu einem globalen Zentrum für Design und Fertigung werden), aber auch die Erhöhung der Limits für ausländische Direktinvestitionen. Indien ist sich der Tatsache bewusst, dass sich die Entwicklung des Landes nur mit Technologie und Innovation vorantreiben lässt, dass ein Verharren in der analogen Welt zum Stillstand führen würde und dass man heute die Weichen stellen muss, um jene Herausforderungen zu meistern, mit denen das Land in einigen Jahrzehnten konfrontiert sein wird. Digitalisierung ist das Credo, mit dem man Indien an die Weltspitze bringen will.

Der Niederösterreicher ist Journalist, Start-up-Experte, Online-Allrounder und Autor zahlreicher Bücher; bekannt sind seine Bestseller *Die Google-Falle* und *Im Visier der Datenjäger*. Gerald Reischl war Redakteur bei den renommierten Tageszeitungen *Die Presse* und *Kurier* und baute das Portal *futurezone.at* zur führenden Technologie-News-Plattform Österreichs auf. Er ist zweifacher Preisträger des New Media Award und wurde 2013 »Onliner des Jahres« (Kategorie Medienmacher/Werbeplanung). 2016 war er Geschäftsführer der ORF Mediaservice GmbH und kümmerte sich um die Start-up-Agenden. Heute ist Reischl als Journalist, Tech-, Trend- und Start-up-Scout sowie Business-Angel aktiv.

The Lower Austrian is a journalist, startup expert, online all-rounder, and the author of numerous books; his best-sellers *Die Google-Falle (The Google Trap)* and *Im Visier der Datenjäger (In the Data Hunters' Sights)* are well known. Gerald Reischl worked as an editor for the prestigious daily newspapers *Die Presse* and *Kurier* and turned the portal *futurezone.at* into Austria's leading technology news platform. He is a two-time winner of the New Media Award and was awarded the title "Onliner of the Year" (in the category Media Maker/Publicity Planning) in 2013. In 2016 he was CEO of ORF Mediaservice GmbH and responsible for the startup sector. Today, Reichl works a journalist and business angel scouting for technologies, trends, and startups.

Gerald Reischl

Digital India: Technology as a Means to Become a World Power Again

In 2060, India will not only be populated by 1.7 billion people: the country is supposed to have turned into the greatest economic power and knowledge society in the world (again) by then. Launching a hitherto unknown digitization campaign, the government intends to provide the infrastructure and conditions necessary to achieve this goal. To do so, it relies on the innovative power of the country's successful companies, its population's will for change and creative drive, and on future issues such a big data, blockchains, artificial intelligence, and the Internet of Things.

India wants to return to where it was three hundred years ago. According to studies by the British economist Angus Maddison, who died in 2010, India and China were the greatest economic powers in the world until the eighteenth century, when both nations were outstripped by Europe in the beginnings of the nineteenth century and then, around the middle of the nineteenth century, by the United States, too. Things have been constantly looking up again for China and India since the turn of the millennium. China's increasingly important position within the worldwide system is well known—India not only wants to follow suit but even top its big neighbor.

Becoming stronger than the rival China

India's present economic growth is remarkable; its gross domestic product (GDP) has risen by 6.75 percent in 2017/18. The Indian government aims at a GDP increase by 7.0 to 7.5 percent for the current fiscal year that ends in March 2019—0.5 to 0.7 percent more than its rival China. This makes India one of the fastest growing economic powers. The success is owed to such initiatives as "Digital India" or "Make in India" (which is supposed to turn India into a global center for design and manufacture), but also to the raised limits for foreign direct investments. Indian players are aware of the fact that the development of their country depends on technology and innovation, that remaining within an analogous world would result in a standstill, and that things have to be put on the right track today in order to rise up to the challenges the country will be confronted with in some decades. Digitization is the creed that is to make India number one in the world.

Indien hat, genauer betrachtet, gar keine andere Wahl. Denn mit 1,3 Milliarden Einwohnern ist es eines der bevölkerungsreichsten Länder der Erde. Laut Schätzung der Vereinten Nationen werden im Jahr 2060 mehr als 1,7 Milliarden Menschen in Indien leben. Das sind rund 18 Prozent der Weltbevölkerung. Die indische Regierung hat erkannt, dass sie jetzt digitale Entscheidungen treffen muss, um dem Land und der Bevölkerung eine Zukunft bieten zu können.

Am 18. Mai 2006 wurde in Indien der nationale E-Governance-Plan mit dem Ziel ausgerollt, allen Indern elektronische Behördenwege zu ermöglichen – zu einer Zeit, in der das Thema Digitalisierung schon im Raum stand, allerdings noch nicht mit Vehemenz forciert wurde. Obwohl der große Erfolg des E-Governance-Plans ausblieb, wurde damit die entscheidende Basis für jenes Programm geschaffen, das Premierminister Narendra Modi neun Jahre später ins Leben gerufen hat: die Initiative »Digital India«. Die Idee kam zur richtigen Zeit: Mitten in den weltweiten Digitalisierungshype hinein wurde die Kampagne am 1. Juli 2015 mit dem Ziel gestartet, Indien in ein total digitales Land zu verwandeln. Modis Vision: Indien müsse zur führenden digitalen Informations-, Wissensgesellschaft und Wirtschaftsmacht werden.

Noch gibt es einen Widerspruch...

Wer heute als Tourist durch Mumbai, New Delhi, Kalkutta oder Varanasi fährt, kann sich kaum ausmalen, dass das 1,3-Milliarden-Land im Jahr 2022 ganz vorne auf dem internationalen Parkett der Digitalisierung mitspielen wird. Indien hinterlässt abseits der indischen Hightech-Hochburg Bangalore einen durch und durch analogen Eindruck. Das beginnt bei der sichtbaren Armut, geht über die Straßenverkäufer, die die Bedarfsartikel des Alltags auf Planen oder Decken auf dem Boden verkaufen, und endet bei jenen Männern, die Touristen mit ihren Fahrradrickschas durch die Städte kutschieren. Viele von ihnen haben schon ein Handy, aber Smartphone-Nutzer sind eher kein Maßstab für den Digitalisierungsgrad einer ganzen Nation.

Wir leben in einer total vernetzten Welt. Zumindest glauben wir das. Smartphones, Smartwatches, Wearables, Tablets, der gegenwärtige Hype um Elektromobilität und selbstfahrende Fahrzeuge, Bitcoins, Blockchain, sprachgesteuerte Assistenten wie Alexa oder Siri machen uns glauben, dass wir in der digitalen Ära angekommen sind und dass das Analoge die Sprache und Welt der Großeltern und Eltern war. Doch es geht noch mehr. Noch viel mehr, denn wir stehen erst am Anfang der digitalen Informationsgesellschaft. 99 Prozent der Welt sind noch nicht vernetzt.

400 Millionen Menschen erhalten digitale Fähigkeiten

Vom »Internet der Dinge« (Internet of Things, IoT) erzählen uns die Tech-Konzerne dieser Welt schon lange, genau genommen seit der Jahrtausendwende. Aus dem IoT ist inzwischen längst das »Internet of Everything« geworden; alles, wirklich alles, kann vernetzt, mit Sensoren oder Chips auf die nächste

On closer scrutiny, India does not have a choice. With its 1.3 billion inhabitants, it is one of the most densely populated countries in the world. According to the United Nations' reckoning, more than 1.7 billion people will live in India by 2060. This amounts to approximately 18 percent of the world population. India's government has realized that digital decisions have to be taken now to offer the country and its population a bright future.

On May 18, 2006, the National e-Governance Plan was launched to make all government services available to the citizens of India via electronic media—at a time in which the issue of digitization had already been touched on, yet not seriously tackled. Although the big success of the e-Governance Plan failed to materialize, it provided the decisive foundation for the program Prime Minister Narendra Modi initiated nine years later: “Digital India.” It was the right time for the idea: aimed at transforming India into a completely digital nation, the campaign was launched on July 1, 2015 right in the middle of the worldwide digitization hype. Modi's vision: India is to become the leading digital information, knowledge and economic power.

A contradictory situation

Driving through Mumbai, New Delhi, Kolkata, or Varanasi as a tourist, one will find it difficult to imagine that the 1.3-billion country will play in the big league of international digitization in 2022. Excepting the Indian high-tech stronghold of Bangalore, the impression India conveys is quintessentially analog. Indicators range from the country's visible poverty and its street vendors offering everyday commodities on tarpaulins and blankets spread on the ground to the men pedaling tourists through the cities on their cycle rickshaws. Many of them have a mobile, yet the use of smartphones is not really a yardstick for an entire nation's degree of digitization.

We live in a completely networked world—or believe to do so at least. Smartphones, smartwatches, wearables, tablets, the present hype around electromobility and autonomous vehicles, bitcoins, blockchains, and language-controlled assistants like Alexa or Siri make us believe that we have arrived in the digital age and that the analog was our grandparents' and parents' world. There is more to come, however—much more. This is just the early stages of the digital information society. Ninety-nine percent of the world are not linked yet.

400 million people to be furnished with digital skills

Technology conglomerates of this world have been telling us stories about the Internet of Things (IoT) for quite some time, in fact since the turn of the millennium. IoT has long since become the Internet of Everything; everything, literally everything, can be linked and advanced to the next level by means of sensors and chips. This is exactly the field India wants to

Stufe gestellt werden. Genau hier will Indien künftig mitspielen. Nicht nur mitspielen, sondern gestalten. Premier Modi dürfte etwas schaffen, was seinen Vorgängern nicht gelungen ist: sämtliche Ministerien auf die für das Land notwendige Digitalisierung einzuschwören, internationale Firmen, die sich schon seit vielen Jahren im Land befinden, mit ins Boot zu holen und nicht zuletzt die indischen Vorzeigekonzerne, die global höchst erfolgreich sind, eine wichtige Rolle im Programm »Digital India« spielen zu lassen.

Das ganze Land, knapp 3,3 Millionen Quadratkilometer groß, soll mit modernster Infrastruktur ausgerüstet werden – nicht nur mit Breitbandinternet via Festnetz, sondern mit Mobilfunk auf Basis der neuen 5G-Technologie, die IoT erst möglich macht. Jeder Inder soll ein eigenes Mobiltelefon haben und mit diesem einen Zugang zu allen Systemen und Behörden des offiziellen Indien. Auf diese Art soll/muss auch jeder ans Bildungssystem angeschlossen werden, denn die Basis jeder Digitalisierung ist eine (aus)gebildete Bevölkerung: Die indische Regierung will das Bildungssystem in den kommenden Jahren forcieren und reformieren; E-Learning, Videokonferenzen, Augmented und Virtual Reality sind Technologien, die dabei eingesetzt werden. Zusätzlich gibt es eine Reihe von nichtstaatlichen Bildungsreinrichtungen. Laut Statistik der University Grants Commission zählt man in Indien 789 Universitäten – darunter nicht weniger als 282 Privatus – und etwa 37.200 Colleges. 400 Millionen Menschen sollen bis 2022 mit digitalen Fähigkeiten ausgestattet werden. Das Ziel der Schaffung zusätzlicher IT-Jobs will man durch enge Kooperationen mit nationalen und internationalen Konzernen erreichen. Im Finanzjahr 2016/2017 flossen etwa 7 Prozent des BIP in soziale Dienstleistungen – davon 2,9 Prozent des BIP in Bildung und 1,4 Prozent in Gesundheit.

Ola und Jio – so funktioniert Innovation in Indien

Statistisch gesehen haben gegenwärtig nur zwei von zehn indischen Haushalten einen Internetzugang, und nur vier von zehn Einwohnern ab 15 Jahren besitzen ein Mobiltelefon. Wobei es auch hier ein Arm-Reich-Gefälle gibt, denn die Mittel- und Oberschicht ist mit den neuesten technologischen Geräten bestens vernetzt und nutzt bereits digitale Angebote. Mit »Digital India« soll diese Kluft geringer werden. Wie schnell etwas in Indien von der Bevölkerung angenommen wird, hat der indische Mobilfunkbetreiber Reliance Jio gezeigt. Erst im September 2016 als reiner 4G-Provider auf Basis der Mobilfunktechnologie LTE (Long-Term Evolution) gestartet, wurde schon im Jahr darauf ein Rekord aufgestellt: Der Datenverbrauch war so hoch wie in keinem anderen Land der Welt. Das mobile Breitbandangebot wurde von der Bevölkerung sofort angenommen, weil Jio Breitbandinternet für jeden erschwinglich machte (ein Gigabyte an Daten kostet umgerechnet 35 Cent). Pro Sekunde melden sich – so gab Mukesh Ambani, Vorstandsvorsitzender von Reliance Industries, Ende Februar 2018 bekannt – sieben Inder für den Mobilfunkservice an. Das dürfte Weltrekord sein.

Während hierzulande über Uber diskutiert wird, Taxifahrer auf die Straße gehen und gegen Innovation demonstrieren, weil sie Erneuerung in der eigenen Branche nicht zulassen, zeigt Indien exakt in diesem Bereich vor, wie Disruption funktioniert. Der indische Uber-Dienst heißt Ola und ist mit 300.000 Fahr-

get involved in, not only by joining the game but by designing it. Prime Minister Modi might achieve something his predecessors were not able to: commit all government departments to the digitization necessary for the whole country; get those international companies on board that have been operating in India for many years; and let India's internationally highly successful model conglomerates play a crucial role in the implementation of the "Digital India" program.

The whole nation with an area of just under 3.3 million square kilometers is to be furnished with state-of-the-art infrastructure: not only with broadband Internet through the landline network but with mobile telephony based on the 5G technology which is *the* requirement for IoT. Every Indian will have a mobile of her or his own and, thus, be able to access all official Indian systems and authorities. This step will/must also connect all Indians to the country's education system since digitization requires a trained/educated population. The Indian government intends to enforce and reform the education system in coming years; e-learning, video conferences, augmented reality, and virtual reality are the technologies the government intends to employ. There is also a number of non-governmental educational institutions in the country. According to a statistical survey by the University Grants Commission, India has 789 universities, including not fewer than 282 private universities, and about 37,200 colleges. 400 million people are to be furnished with digital skills by 2022. Close collaboration with national and international conglomerates is supposed to ensure the creation of additional IT jobs. 7 percent of the GDP were spent on social services in the fiscal year 2016/17—2.9 percent of the GDP on education and 1.4 percent on health.

Ola und Jio: this is how innovation works in India

Statistics show that today only two out of ten Indian households have access to the Internet; four out of ten inhabitants older than fifteen own a mobile. There is a gap between the rich and the poor also in this regard; employing the most recent technology devices, the middle and upper classes are perfectly connected and already benefit from digital offers. "Digital India" aims at narrowing this gap. The Indian mobile network operator Reliance Jio has shown how quickly things may be accepted by the population. Launching its 4G services based on the mobile telephony standard LTE (Long-Term Evolution) only in September 2016, the company already set a record in the year after: the data consumption was higher than in all other countries of the world. The mobile broadband offer was immediately taken up by the population since Jio's broadband Internet was affordable for everybody (the tariff for one gigabyte is 35 cent). Mukesh Ambani, CEO of Reliance Industries, announced in late February 2018 that seven people register for the mobile telephony service every second. This seems to be a world record.

Whereas Uber has become an issue of discussion here, with taxi drivers taking to the street to protest against innovation and rejecting a transformation of their business, India demonstrates how disruption

zeugen die populärste Taxi-, Sammeltaxi- und Mietwagen-Plattform des Landes. Gegen dieses junge Unternehmen, das ebenfalls kein eigenes Auto besitzt, sondern »nur« vermittelt, wird – im Gegensatz zu Europa – nicht gesetzlich eingeschritten. Ein Viertel der Bevölkerung nutzt Ola, das in mehr als hundert indischen Städten angeboten wird, und bezahlt diese Dienstleistung bargeldlos via App.

Die Infrastruktur muss modernisiert werden

Die Regierung investiert in den Straßenausbau, in das Eisenbahnnetz und in moderne Häfen und Flughäfen. Urbane Zentren müssen mit Infrastruktur für die Energie- und Wasserversorgung ausgestattet werden. Ein Monstervorhaben, denn anders als in Europa, wo es analoge Kataster gibt, aus denen sich Wasserleitungen, Kanalverläufe und Stromleitungen herauslesen lassen, fehlen diese Daten in den meisten Städten Indiens und müssen erst aufwendig erstellt werden. Zieht man den Mobilfunk als Beispiel heran, kann man davon ausgehen, dass das schnell geschehen kann – beziehungsweise geschehen muss. Laut einer Studie von McKinsey braucht Indien pro Bewohner 140 Liter Wasser pro Tag, kann aber nur 105 Liter liefern. Um nicht in eine ähnliche Situation wie Südafrika zu kommen, wo Wasser bereits rationiert werden muss, setzt man mithilfe von us-Konzernen wie etwa Cisco darauf, den gesamten Bereich Wasser, vom Trinkwasser bis zum Abwasser, zu digitalisieren. In Andhra Pradesh hat Cisco eine Wassermanagementlösung auf Basis von IoT geschaffen, um eine drohende Wasserkrise zu verhindern.

Einen weiteren guten Grund, warum Indien gar nichts anderes übrigbleibt, als die digitale Offensive voranzutreiben, liefert auch die Statistik der indischen Energiebehörde Central Electricity Authority (CEA): 300 Millionen Inder leben nach wie vor ohne Stromversorgung. Aber bereits 2022 soll jeder Bürger an das Stromnetz angeschlossen sein.

Smart Grid und künstliche Intelligenz

Wer heute von Energie spricht, muss im selben Atemzug auch Smart Grid, Big Data und künstliche Intelligenz (KI) nennen – Schlagworte, mit denen Technologiekonzerne und Regierungen auf der ganzen Welt die Zukunft in neue Begriffe fassen. Das sind auch in Indien jene Technologien, mit denen das Land auf Innovationskurs gebracht werden soll. Strom- beziehungsweise Energienetze sollen smart werden: Die einzelnen Akteure im Energiesystem sind über ein Kommunikationsnetzwerk miteinander verbunden und sorgen für ein intelligentes, völlig flexibles Energienetz, das Strom in ausreichender Menge genau dann liefert, wenn er gebraucht wird. Es werden neue Geschäftsmodelle erstellt, und in Echtzeit weiß man, wie es um das Netz steht: Laufend werden digital Daten gesammelt und ausgewertet. Diese Daten liefern nicht nur wichtige Informationen über die optimale Stromverteilung, sondern auch über die ideale Stromgenerierung. Hier schließt sich wiederum der Kreis zu den erneuerbaren Energien, denn mit künstlicher Intelligenz werden die Echtzeitdaten mit aktuellen Wetterdaten und Prognosen abgestimmt, um den künftigen Energiebedarf abschätzen zu können.

works in this very field. The Indian Uber service, called Ola, boasts 300,000 vehicles and is the most popular taxi, share taxi, and rental car platform in the country. Contrary to Europe, no legal action is taken against the young company that does not own any cars either but merely provides rides. A quarter of the population uses Ola, which operates in more than one hundred Indian cities, and pays for its services cashless via an app.

India's infrastructure calls for modernization

India's government invests in road improvement, in the rail network, as well as in modern harbors and airports. Urban centers have to be furnished with energy and water supply infrastructure—an enormous undertaking considering that unlike European cities, where analog cadasters contain information on water pipes, sewer systems, and power lines, most municipalities of India lack such data; it requires great efforts to provide the relevant facts and figures. Taking mobile telephony as an example, one may assume that this can and must be done promptly. According to a study by McKinsey, India's water demand amounts to 140 liters per person a day; yet only 105 liters can be supplied. To avoid a situation like that in South Africa, where water has to be rationed, the Indian government relies on the support of us conglomerates like Cisco to digitize the whole range of aspects from drinking water to waste water. Cisco already operates an IoT-based water management facility in Andhra Pradesh, which is supposed to prevent a threatening water crisis.

A statistical survey undertaken by India's Central Electricity Authority (CEA) provides the responsible Indian bodies with another good reason for pressing ahead with the digital campaign as their only option: 300 million Indians still live without electricity. However, all citizens are supposed to be supplied with electricity by 2022.

Smart grid and artificial intelligence

Speaking about energy today, one cannot but mention smart grids, big data, and artificial intelligence (AI) in the same breath. Technology conglomerates and governments all over the world resort to these catchwords to redefine tomorrow's reality. These technologies are also those India banks on for putting the country on a track of innovation. The authorities plan to turn electricity and other energy grids into smart grids, to connect all the energy systems' individual players within a communication network in order to ensure an intelligent, completely flexible energy grid that supplies sufficient amounts of energy exactly when they are needed. New business models are developed, and real-time information on the operation of the grid can be provided. Digital data are continuously gathered and evaluated. These data offer important details not only on the best possible distribution but also on the best possible generation of electricity. This ties in with the issue of renewable energies, with real-time data being matched with the latest weather data and forecasts by means of AI in order to assess future energy demands.

Es klingt beinahe utopisch, dass Indien in diesem Bereich ein globaler Vorreiter werden will – man bedenke nochmals die Zeitachse: bis 2022. Aber es gibt bereits internationale Beispiele, wie andere Staaten, die technologisch völlig hinterherhinkten, innerhalb weniger Jahre zu internationalen Musterschülern und Vorreitern wurden. In Europa bieten die baltischen Staaten ein Musterbeispiel für eine solch rasante Entwicklung. Relativ spät in die Mobilfunkära eingestiegen, gehören sie jetzt nicht nur zu den modernsten Mobilfunknationen der Welt, sondern sind auch jene, die das Thema E-Government (mittlerweile auch oft als M-Government) bezeichnet, besetzt haben. Ein ähnliches Bravourstück könnte auch Indien gelingen, da dank der Initiative »Digital India« alle Kräfte im Land an einem Strang ziehen und an der Realisierung der von der Regierung vorgegebenen Ziele arbeiten.

Hinzu kommt das gesamte Thema Wohnraum, das in den kommenden Jahren eine große Herausforderung sein wird. Bereits jetzt leben 30 Prozent der 1,3 Milliarden Einwohner Indiens in Städten, im Jahr 2030 sollen es schon 600 Millionen sein. Diese Zuwanderung stellt die lokalen Verwaltungen vor große Aufgaben, denn schon heute können die Bewohner der Städte nicht ausreichend mit sauberem Wasser, sauberer Luft und Energie versorgt werden. Tatsache ist, dass Indien diese Vorhaben alleine nicht schaffen wird und dass Unternehmen aus dem Ausland bei diesen technisch äußerst anspruchsvollen Projekten gefragt sind.

E-Government für alle Bürger

E-Government soll in einigen Jahren zum indischen Alltag gehören, und dazu zählen nicht nur Amtswege, die man digital erledigen kann. Auch die digitale Identität – bei uns in Europa noch vieldiskutiert – wird Teil des indischen Alltags sein. Mit dieser elektronischen ID soll der Bevölkerung der Zugang zu einem neuen (Wirtschafts-)Leben geöffnet werden: angefangen beim Bildungssystem, in das man mittels E-ID Zugang erhält (vom Offline- bis zum Onlinekurs), über Onlinebanking (fast 300 Millionen Inder besitzen übrigens noch gar kein Konto) bis zu Behördenwegen, die ganz ins Web verlagert werden sollen – auch Betriebsgenehmigungen, Zertifikate und dergleichen sollen online abrufbar sein. Die Digitalisierung Indiens wird von vielen indischen Unternehmen und Banken mitgetragen. Die Bandhan Bank etwa hat 2014 als Microbanking-Institut begonnen, um arme Menschen ins Finanzsystem zu integrieren. Mittlerweile gibt es etwa 2700 Minifilialen im Land, und – besonders beeindruckend – 16 Millionen Menschen nutzen bereits das von der Bank angebotene E-Payment-System, bei dem Kunden bargeldlos via Bandhan-Bank-App bezahlen können.

IndiaChain wird größtes Blockchain-Netzwerk

Indien hat einen eigenen, von der Regierung ins Leben gerufenen Thinktank, NITI (National Institution for Transforming India) Aayog, der sich mit Nachhaltigkeitszielen beschäftigt und den Regierungsmitgliedern, allen voran Premierminister Narendra Modi, wichtigen Input gibt und damit Entscheidungs-

Considering the time horizon until 2022, it may appear to be an almost utopian notion that India wants to become a global pioneer in this field. Some international examples, however, show that other states lagging miles behind technologically have turned into international models and trailblazers within only a few years. In Europe, the Baltic States offer a classic example of such a rapid development. Though having entered the era of mobile telephony at a relatively late point in time, these states not only rank among the most advanced mobile telephony nations in the world today but have also made e-government (also called mGovernment in the meantime) their concern. India might bring off a similarly daring feat as, thanks to the “Digital India” initiative, all players in the country have joined forces to achieve the goals set by the government.

Another crucial issue is housing, which will present a considerable challenge in the years to come. Thirty percent of India’s 1.3 billion inhabitants live in cities today, and their number will have risen to 600 million by 2030. Local administrations will find themselves charged with a number of difficult tasks because of this immigration: even the inhabitants of today’s Indian cities cannot be supplied with sufficient clean water, clean air and energy. There can be no doubt that India will not be able to tackle these challenges by itself and that foreign companies will have to get involved in this technologically very demanding projects.

E-government for all citizens

E-government is to become an integral part of India’s everyday world, and this does not only include the digital delivery of services in the public sector. Digital identity—a still very controversial subject in Europe—will be a day-to-day reality. This electronic form of identification is to make a new (economic) life accessible to the population. The measures span from the education system—for which a special identification will facilitate access (to an offer ranging from offline to online courses)—and online banking (300 million Indians have no bank account yet, by the way) to administrative procedures that will all be moved into the Internet (operating licenses, certificates, and many other things will be also retrievable online). India’s digitization is supported by many of the country’s companies and banks. The Bandhan Bank, for example, has started out as a microfinance institution in 2014, with the intention of integrating poor people into the financial system. Today, there are 2,700 mini-branches in the country, and—which is particularly impressive—16 million clients use the e-payment system the bank offers, which enables them to pay cashless via the Bandhan Bank app.

IndiaChain to become largest blockchain network

India has its own think tank, NITI (National Institution for Transforming India) Aayog, which was established by the government. NITI Aayog explores sustainable development goals and provides important input to members of the government, especially to Prime Minister Narendra Modi, offering a basis for future decisions. The think tank is also responsible for the major blockchain campaign and works on estab-

grundlagen für die Zukunft liefert. NITI Aayog ist auch für die große Blockchain-Offensive verantwortlich und arbeitet daran, mit IndiaChain das größte Blockchain-Netzwerk des Landes aufzubauen. Eine »Blockkette« ist eine Datenbank, an deren »Anfangsblock« nach einem Check chronologisch weitere Blöcke angehängt und verschlüsselt werden, ohne dass Blöcke davor gelöscht oder bearbeitet werden können. Dieses Datenbanksystem soll künftig in verschiedensten Bereichen angewendet werden. In Indien soll die Technologie unter anderem auch dabei helfen, Onlinebetrug zu verhindern, Vertragsunterzeichnungen zu beschleunigen, Transaktionen transparenter zu machen und die Landwirtschaft zu reformieren.

Dass die Blockchain-Technologie für Verifikation und gegen Onlinebetrug genutzt werden kann, ist bekannt, der Einsatz in der Landwirtschaft hingegen ist neu. Es sollen digitale Akten für Ackerland entstehen, in denen man quasi die Identität jedes Quadratmeters feststellen wird können – was darauf gepflanzt wurde und wie die Beschaffenheit des Bodens aussieht, welche Mineralien sich im Boden finden, etc. Die Änderungen dieser Bodendaten sollen mithilfe der Blockchain-Technologie beobachtet werden.

Generell wird in Indien großes Augenmerk auf Cybersicherheit gelegt. Man weiß: Die smarte Welt muss sicher sein.

Technologie-Hotspot Bangalore

Bangalore ist als der Technologie-Hotspot Indiens bekannt, gilt als Zentrum der Luft- und Raumfahrtindustrie und ist Forschungshochburg: Es gibt 400 Forschungszentren in der Stadt, jeder zweite Forscher arbeitet in Bangalore, so gut wie jeder bekannte IT-Konzern hat in der 12-Millionen-Metropole eine Niederlassung. Österreiche Unternehmen wie etwa Austrian Airlines oder die Erste Bank lassen seit Jahren gewisse IT-Leistungen von dort ansässigen Unternehmen erledigen. 4500 IT-Firmen zählt man in Bangalore, 2 Millionen Menschen arbeiten direkt in diesen Unternehmen. Rechnet man die Zulieferer dazu, sind es gar 4,5 Millionen Menschen..

Jene allgemein als innovativ verstandenen Technologien, mit denen man Probleme rund um den Globus lösen will – KI und Big Data –, stehen derzeit wie eine große Klammer über allen Herausforderungen der Zukunft. Kein Wunder also, dass auch bei der jüngst ausgerufenen Start-up-Offensive exakt jene Jungunternehmen, die sich mit Lösungen in diesem Bereich beschäftigen, als besonders nachhaltig, investitions- und förderungswürdig eingestuft werden. Indien soll ein KI-Powerhouse werden. Am 1. Februar 2018 hat Indiens Finanzminister Arun Jaitley ein staatliches Programm für künstliche Intelligenz angekündigt. Allein für 2018 wurden im Budget Mittel in Höhe von umgerechnet fast 400 Millionen Euro für Lösungen in den Bereichen 3-D-Printing, künstliche Intelligenz und Industrie 4.0 vorgesehen.

lishing IndiaChain and make it the largest blockwork chain in the country. A blockchain is a continuously growing concatenation of blocks, which, once recorded, cannot be deleted or altered retroactively. The government plans to use this data base system in a wide variety of fields. The technology to be implemented in India is to help to prevent Internet fraud, speed up the signing of contracts, make transactions more transparent, and reform the agricultural sector.

That the blockchain technology can be used for verification processes and as an instrument against Internet fraud is known; relying on it for agricultural issues is something new, however. The digital arable land register to be compiled will encompass the identity of every square meter, as it were—what has been planted there, the properties of its soil, which minerals the area contains, etc. The blockchain technology will monitor these soil data and how they change.

Indian authorities generally pay careful attention to cybersecurity. They know that the smart world must be secure.

Technology hotspot Bangalore

Bangalore has come to be known as India's technology hotspot. It is considered the center of aviation and space travel industry and a science stronghold. The city boasts 400 research institutions; every other scientist works in Bangalore, and virtually every IT conglomerate of distinction has a subsidiary in the twelve-million-metropolis. Austrian companies such as Austrian Airlines or Erste Bank have outsourced certain IT tasks to Bangalore-based firms for years. 2 million people work directly in the city's 4,500 IT companies. Including their outside suppliers, the number of employees amounts to 4.5 million.

The technologies generally regarded as innovative that are expected to solve problems across the world—AI and big data—form a huge bracket encompassing all future challenges. This is why it came as no surprise that the government's recent announcement of a campaign for start-ups categorized exactly those fledgling business enterprises focusing on solutions in this field as particularly sustainable and worth supporting and investing in. India is to be turned into an AI powerhouse. On February 1, 2018, India's Finance Minister and Minister of Corporate Affairs Arun Jaitley presented a government program for artificial intelligence. The budget provides for the equivalent of almost 400 million euros for solutions in the fields of 3D printing, artificial intelligence, and Industry 4.0 just for 2018.

India's Silicon Valley—start-up nation

"We want to become India's Silicon Valley," says Pankajam Sridevi, president of Bangalore's Chamber of Industry and Commerce. We have (too) often heard such statements from other parts of the world in the meantime. Fact is, however, that the start-up campaign launched from Bangalore is not just supposed to cover the entire nation but also allow India to catch up with China and join the players in the international start-up arena.

Das Silicon Valley Indiens – Start-up-Nation

»Wir wollen das Silicon Valley Indiens werden«, sagt die Präsidentin der Industriellenvereinigung Bangalores, Pankajam Sridevi. Diesen Satz hat man mittlerweile auch aus anderen Teilen der Welt schon (zu) oft gehört. Tatsache ist aber, dass von Bangalore aus eine Start-up-Initiative aufgesetzt wurde, die in den kommenden Jahren nicht nur ganz Indien erfassen soll, sondern mit der Indien zu China aufschließen und auf dem internationalen Start-up-Parkett mitspielen will.

Im internationalen Ranking der besten Start-up-Ökosysteme rangiert Bangalore auf dem 20. Platz, gilt nach Berlin als das am schnellsten wachsende Start-up-System und hat mit einem Durchschnittsalter von 29 Jahren die jüngsten Unternehmer; im Silicon Valley beträgt das Durchschnittsalter 36 Jahre. Dort zählt man übrigens 120.000 Start-ups; pro Tag werden im Valley statistisch gesehen 47 neue Start-ups gegründet. Der us-Konzern Cisco, der in Indien bereits mit einigen IoT-Projekten wie etwa dem erwähnten Wassermanagementsystem präsent ist, hat 280 Millionen Dollar für Investitionen in indische Start-ups reserviert – an gut zwei Dutzend Start-ups ist Cisco bereits beteiligt.

HACKccelerator

Die Start-up-Welle soll von Bangalore ausgehen, wo es mittlerweile nicht weniger als 140 Akzeleratoren gibt. Der größte heißt »10.000 Start-ups« und wurde vor vier Jahren von der Vereinigung von IT-Unternehmen »Nasscom« gestartet. HACKccelerator hat mittlerweile Niederlassungen in zehn indischen Städten. Das Start-up-Warehouse verfolgt das Ziel, in den kommenden zehn Jahren 10.000 indische Tech-Start-ups zu ermitteln. Auch wenn es angesichts von 1,3 Milliarden Indern ein Leichtes sein sollte, dieses Ziel zu erreichen, gibt es laut Angaben von »10.000 Start-ups«-Direktorin Kritika M. allerdings derzeit im ganzen Land nur 7700 Tech-Start-ups.

Das »10.000 Start-ups«-Programm läuft seit 2013. 5000 Start-ups haben das Programm bereits durchlaufen, für 327 fand man Investoren. »Uns ist aufgefallen, dass die Start-up-Gründer früher zwischen 23 und 26 Jahre alt waren«, sagt Kritika M. »Inzwischen beträgt das Alter 38 bis 54 Jahre.« Die Start-ups lernen also, dass es sinnvoll ist, Ältere und Erfahrene im Team zu haben, die mit ihrem Netzwerk punkten und die bereits Vertrauen bei den Investoren genießen. Kritika M.: »Mit einem Älteren an Bord kommt man leichter zu einem Funding.« Für das Start-up-Programm bewirbt man sich online; eine Jury entscheidet, ob man zu einem der monatlich stattfindenden Demo Days eingeladen wird. Bis zu 15 Start-ups werden pro Monat in das Programm, das ein Jahr dauert, aufgenommen. »Wir nehmen keine Anteile, aber jedes Start-up muss pro Mitarbeiter 60 Dollar im Monat zahlen«, sagt Kritika M. Maximal zwölf Mitarbeiter dürfen sich im Start-up-Warehouse einquartieren.

Die Start-up-Aktivitäten der Konzerne – Cooler als Google

Parallel zur Tätigkeit der offiziellen Akzeleratoren, die rein auf indische Start-ups abzielen, gibt es auch einige Konzerne mit eigenen Start-up-Initiativen. So etwa den bekannten Mischkonzern Tata (Tata Consultancy Services). In den Bereichen Gesundheit, Telekommunikation, Auto (Tata ist unter anderem

The international start-up ecosystem ranking report lists Bangalore as number 20. The city is considered the fastest growing start-up system after Berlin and boasts the youngest entrepreneurs with an average age of twenty-nine years compared to thirty-six years in the Silicon Valley with its 120,000 start-ups; statistically, forty-seven start-ups are founded in the Valley each day. The us technology conglomerate Cisco, which is already present in India with some IoT projects such as the aforementioned water management system, has allocated 280 million dollars for investments in Indian start-ups; Cisco already holds a stake in two dozen start-ups.

HACKccelerator

The start-up wave is supposed to start from Bangalore with its not fewer than 140 accelerators. The biggest is called “10,000 Start-ups” and was launched by the association of IT companies Nasscom. In the meantime, HACKccelerator has subsidiaries in ten Indian cities. The start-up warehouse intends to find 10,000 Indian tech start-ups within the following ten years. Even if—considering India’s population of 1.3 billion citizens—this goal seems easy to achieve, there are only 7,700 tech start-ups all across India today according to Kritika M, director of the initiative “10,000 Start-ups.”

The program “10,000 Start-ups” has been in operation since 2013. 5,000 start-ups have already passed through the program; investors were found in 327 cases. “We realized that start-up founders were between twenty-three and twenty-six years old in the past,” says Kritika M. “Today, their age lies between thirty-eight and fifty-four years.” Start-ups have obviously learned that it makes sense to have older and experienced people in the team who score with their networks and already enjoy the investors’ confidence. Kritika M: “It is easier to get funding if you have an older person on board.” Aspirants apply online for the start-up program; a jury decides whether an applicant will be invited to one of the monthly demonstration days. Each month, up to fifteen start-ups are accepted into the one-year program. “We do not get a share, but every start-up has to pay sixty dollars per person and month,” says Kritika M. Not more than twelve employees may take up quarters in the start-up warehouse.

The conglomerates’ start-up activities: cooler than Google

Parallel to the official accelerators’ activities exclusively aimed at Indian start-ups, some players such as the well-known mixed conglomerate Tata (Tata Consultancy Services), which has 600,000 employees, pursue their own start-up initiatives. Tata collaborates with start-ups—that were discovered either at relevant international events like “Slush” in Finland or directly contacted Tata on their own initiative—in the fields of health, telecommunication, automobiles (Tata owns Jaguar and Landrover, among others), food, robotics, drones, etc. The conglomerate Infosys, which employs 200,000 people around the world, also cooperates with international start-ups and intends to intensify its search for young companies. Infosys is probably one of the best examples of how inspiring successful conglomerates of the tech scene can be. The company head-

Eigentümer von Jaguar und Landrover), Nahrungsmittel, Robotik, Dronen und so weiter kooperiert der 600.000-Mitarbeiter-Konzern mit Start-ups, die entweder bei einschlägigen internationalen Events wie etwa dem finnischen »Slush« entdeckt wurden oder sich direkt beim Unternehmen gemeldet haben. Auch der Konzern Infosys – 200.000 Mitarbeiter weltweit – kooperiert mit internationalen Start-ups und will die Suche nach jungen Unternehmen forcieren. Infosys ist vermutlich eines der besten Beispiele, wie sehr man sich von den erfolgreichen Konzernen der Tech-Szene inspirieren lässt. In Bangalore hat Infosys eine Firmenzentrale aufgebaut, die ebenso gut in Kalifornien stehen könnte: ein großer, bunter Campus für die Mitarbeiter, mit einem Coolness-Faktor, wie man ihn auch bei Google & Co im Silicon Valley spürt. Nicht nur Restaurants, ein Kino, ein Waschsalon, Fitnesscenter und Supermärkte sind hier zu finden, es gibt sogar ein eigenes Hotel, damit Mitarbeiter von anderen Firmenstandorten direkt am Campus übernachten können.

Auch Österreich engagiert sich in Indien

»Indien sucht den Anschluss an das internationale Start-up-Parkett«, sagt Raj Kumar Srivastava vom Forschungsrat Karnatakas. »Indien ist der am schnellsten wachsende Markt und wird China demnächst überholen«, sagt der Vorsitzende des Rats für Forschung und Technologieentwicklung (RFT), Hannes Androsch, der als Industrieller schon seit knapp 20 Jahren mit einem Werk von AT&S in Indien engagiert ist. »Indien ist ein aufstrebendes Land, Teil der globalen Bedeutungsverschiebung weg vom Atlantik hin zum indopazifischen Raum. Es wird in den kommenden Jahren Wachstumsraten von sieben bis acht Prozent geben. Man muss rechtzeitig da sein und die Möglichkeiten nutzen.«

Da auch österreichische Forscher und Start-ups vom Aufstieg Indiens profitieren sollen, ist eine engere wissenschaftliche Zusammenarbeit zwischen Österreich und Indien geplant. Das Austrian Institute of Technology (AIT), das Indian Institute of Science und die Außenwirtschaft Österreich bereiten eine groß angelegte Kooperation in Sachen Forschung vor. Wie diese konkret funktionieren kann, soll in zwei Workshops in den beiden Ländern geklärt werden. »Man braucht hinreichend viele Gemeinsamkeiten, muss aber gleichzeitig komplementär aufgestellt sein, um eine Win-Win-Situation zu erzeugen«, sagt AIT-Chef Wolfgang Knoll. Organisatorisch soll das »Zwei-plus-zwei-Modell« der Max-Planck-Gesellschaft Vorbild sein, bei dem nicht nur zwei wissenschaftliche Institutionen kooperieren, sondern von Anfang an auch zwei Unternehmen mit an Bord sind. Knoll sieht hier das AIT als »Türöffner für österreichische Unternehmen«. ✘

quarters Infosys built in Bangalore might as well be in California: a large, colorful campus for the employees with a coolness factor that can also be felt at Google's and other Silicon Valley companies' HQs. You will not only find restaurants, a cinema, a laundromat, a fitness center, and supermarkets but also a special hotel so that employees from other locations of the company can spend the night on campus.

Austria's involvement in India

"India strives to join the international start-up scene," says Raj Kumar Srivastava, who works for the Research Board of Karnataka. "India is the fastest growing market and will soon outshine China," says Hannes Androsch, chairman of the Council for Research and Technology Development, who has been involved in India with an AT&S plant as an industrialist for almost twenty years. "India is an emergent country, part of the global shift of significance from the Atlantic to the Indo-Pacific sphere. Growth rates will amount to seven to eight percent. One has to be there in time and take advantage of the possibilities offered."

Since Austrian researchers and start-ups are to benefit from India's rise, too, a closer collaboration between Austria and India has been envisaged. The Austrian Institute of Technology (AIT), the Indian Institute of Science, and International Advantage Austria (Außenwirtschaft Österreich) are developing a program for a large-scale collaboration between the two countries. Two workshops in the two countries will explore the details of this undertaking. "Such a collaboration requires sufficient common ground on the one hand but also complementary interests on the other if a win-win situation is to be created," says AIT managing director Wolfgang Knoll. In terms of organization, the cooperation is to be based on the "two-plus-two model" of the Max-Planck-Gesellschaft, which involves not only two research institutions but also two companies from the start. Knoll regards AIT as "a door opener for Austrian companies" in this case. ✕

Alpbacher Technologiegespräche

»Diversität und Resilienz«

23. – 25.08.2018

Die Alpbacher Technologiegespräche werden von AIT – Austrian Institute of Technology, Österreichs größter Research-and-Technology-Organisation, und ORF Radio Österreich 1 veranstaltet. Das Projekt wird von Mag. Michael H. Hlava (AIT) und Dr. Martin Bernhofer (ORF Ö1) geleitet, das Projektbüro von Claudia Klement (AIT). Dem Steering Committee der Alpbacher Technologiegespräche gehören Dr. Hannes Androsch (Vorsitzender des Aufsichtsrates des AIT, Vorsitzender Forschungsrat RFE), Prof. Dr. Wolfgang Knoll (wissenschaftlicher Geschäftsführer AIT) und Monika Eigensperger (Radiodirektorin ORF) an.

Wissenschaftlicher Partner der Alpbacher Technologiegespräche 2018 ist die Helmholtz-Gemeinschaft Deutscher Forschungszentren, Industrial Partner die Industriellenvereinigung (IV).

Die Veranstaltung wird maßgeblich vom österreichischen Bundesministerium für Verkehr, Innovation und Technologie (BMVIT), vom Bundesministerium für Digitalisierung und Wirtschaftsstandort (BMDW) sowie vom österreichischen Bundesministerium für Bildung, Wissenschaft und Forschung (BMBWF) unterstützt.

Große Innovationen entwickelten sich nie in abgeschotteten Monokulturen. Zentren der Innovation und der Hochkulturen waren immer multikulturell, geprägt durch lebhaften Handel und Städte, die innovative Geister anzogen. Das galt für die frühen Kulturen in China, Indien oder etwa Babylon, das heute noch für die im Alten Testament erwähnte »Sprachverwirrung« bekannt ist, ebenso wie für heutige Innovationsstätten wie das Silicon Valley oder noch viel jüngere, sehr aufstrebende Städte wie Bangalore oder Seattle, die heuer Gast bei den Alpbacher Technologiegesprächen sind. Die Alpbacher Technologiegespräche machen sich auf die Spurensuche, um das Zusammenspiel von Innovation, Diversität und Resilienz zu ergründen. Erst die Durchmischung ermöglicht es, den eigenen Horizont zu erweitern und von der »Welt« zu lernen. Oft war es die mangelnde Diversität, das Entstehen von Monopolen und Monokulturen, die große Kulturen, Städte oder auch Unternehmen zu Fall brachten. Nur Vielfalt schafft – egal ob in der Gesellschaft oder in der Biologie – die nötige Widerstandsfähigkeit von Systemen. Diversität und Resilienz sind wichtige Faktoren für Innovation und Fortschritt. Doch wie sieht es damit in unserer zunehmend digitalisierten Welt aus, in der Big Data, künstliche Intelligenz und virtuelle Welten teils völlig neue Voraussetzungen schaffen? ✕

Alpbach Technology Symposium

“Diversity and Resilience”

23. – 25.08.2018

The Alpbach Technology Symposium is organized by AIT – Austrian Institute of Technology, Austria’s largest research and technology organization, and ORF Radio Österreich 1. The project is managed by Mag. Michael H. Hlava (AIT) and Dr. Martin Bernhofer (ORF Ö1); Claudia Klement (AIT) is head of the project office. The Alpbach Technology Symposium’s steering committee includes Dr. Hannes Androsch (Chairman of the Supervisory Board of AIT, Chairman of the Austrian Council for Research and Technology Development), Prof. Dr. Wolfgang Knoll (Scientific Managing Director of AIT) and Monika Eigensperger (Director of ORF Radio).

Scientific partner of the Technology Symposium 2018 is the Helmholtz Association of German Research Centres, the industrial partner is the Federation of Austrian Industries (IV).

The Symposium is substantially supported by the Austrian Federal Ministry for Transport, Innovation and Technology (BMVIT), the Austrian Federal Ministry for Digital and Economic Affairs (BMDW), and the Austrian Federal Ministry of Education, Science and Research (BMBWF).

Major innovations have never arisen in isolated monocultures. Centers of innovation and advanced civilizations have always been multicultural, characterized by vigorous trade and cities that attracted innovative minds. That was true for the early civilizations in China, India, and Babylon, still known today for the “confusion of tongues” mentioned in the Old Testament, and still holds true for today’s innovation hubs like Silicon Valley or even newer aspiring cities like Bangalore or Seattle. The Alpbach Technology Symposium has set out on a search to explore the interplay of innovation, diversity, and resilience. Only a functioning combination of these factors can make it possible to broaden our own horizon and learn from “the world.” Very often a lack of diversity and the rise of monopolies and monocultures have brought down major cities or enterprises. Whether in society or in biology, only diversity provides systems with the necessary resilience. Diversity and resilience are important factors for innovation and progress. But what does that mean for our increasingly digitized world, where big data, artificial intelligence, and virtual worlds are creating completely different requirements? ✕

Jahrbuch zu den Alpbacher
Technologiegesprächen 2018
Alpbach Technology Symposium
Yearbook 2018

TEC

Alpbach Technology Symposium Alpbacher Technologiegespräche

Idee und Konzept **Idea and concept**

Hannes Androsch, Michael H. Hlava, Martin Kugler
Alpbacher Technologiegespräche

Sarah Hellwagner, Clemens Kopetzky
art: phalanx Kommunikationsagentur GmbH, Wien

Medieninhaber **Media owners**

art: phalanx Kommunikationsagentur GmbH, Wien
Verlag Holzhausen GmbH

Herausgeber **Publishers**

Hannes Androsch, Wolfgang Knoll, Anton Plimon
und art: phalanx Kommunikationsagentur GmbH,
Wien

Produktion **Production**

Sarah Hellwagner, Lisa Reichkendler
art: phalanx Kommunikationsagentur GmbH, Wien

Redaktion **Editor**

Martin Kugler

Lektorat **Copy-editing**

Wolfgang Astelbauer, Wien

Übersetzung **Translations**

Wolfgang Astelbauer, Wien; Michael Strand, Wien;
Brigitte Willinger, Wien

Grafisches Konzept und Gestaltung

Visual concept and design

ap media – Visuelle Kommunikation,
Barbara Hoffmann, Andreas Soller, Wien

Lithografie **Lithography**

Pixelstorm, Wien

Schriften **Fonts**

Vista Sans, Vista Slab (Xavier Dupré)

Papier **Paper**

Munken Polar, Fedrigoni Tatami White

Druck **Printed by**

Medienfabrik Graz GmbH

Verlag **Publishing house**

Verlag Holzhausen GmbH

1. Auflage 2018
First Edition 2018

Bildnachweis **Photo credits**

S./p. 6: © AIC, Peter M. Mayr; S./p. 18: © AIT, Krischanz,
zeiller; S./p. 20: © AIT, Jessica Karle; S./p. 80: © AIT,
Rita Skoff; S./p. 102: © AIT, H. Krischanz; S./p. 112:
© Foto Schoerg; S./p. 146: © Medizinische Universität
Wien; S./p. 164: © Heimo Aga; © Christine Knoll-
Ramach; S./p. 174: © I.E.C.T., Nick Liseiko Photogra-
phy; S./p. 190: © Christian Skalnik

© 2018 Verlag Holzhausen GmbH

© 2018 art: phalanx Kommunikationsagentur
GmbH, Wien

Alle Rechte, auch die des auszugsweisen Abdruckes
oder der Reproduktion einer Abbildung, sind vor-
behalten. Das Werk einschließlich aller seiner Teile
ist urheberrechtlich geschützt. Jede Verwertung
bedarf der Zustimmung der art: phalanx Kommuni-
kationsagentur GmbH, die Verwertung des E-Book
auch der Zustimmung des Holzhausen Verlages.

All rights are reserved, including the rights to
copy extracts or reproduce illustrations. Any
and all parts of this work are protected by copy-
right. Any kind of use requires the consent of
art: phalanx Kommunikationsagentur GmbH,
the use of the e-book the additional approval of
Holzhausen Verlag.

ISBN 978-3-903207-24-0

E-Book ISBN 978-3-903207-25-7

Printed in Austria, EU

Alle Rechte vorbehalten

All rights reserved

Bibliografische Information der Deutschen

Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese
Publikation in der Deutschen Nationalbibliografie;
detaillierte bibliografische Daten sind im Internet
über <http://dnb.dnb.de> abrufbar.

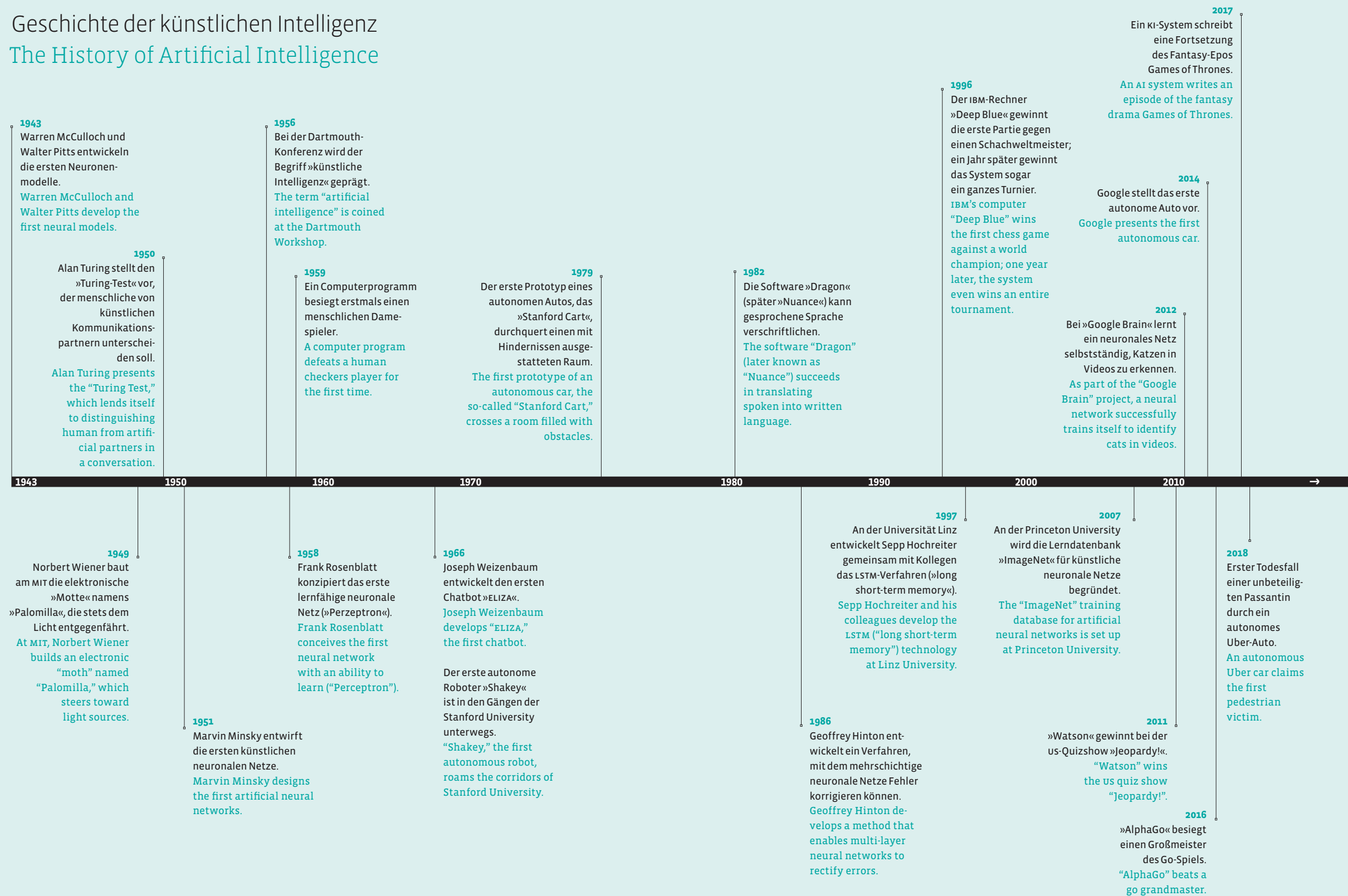
Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publi-
cation in the Deutsche Nationalbibliografie;
detailed bibliographic data are available on the
Internet at <http://dnb.dnb.de>.

Aus Gründen der Lesbarkeit wurde verzichtet, durch-
gängig geschlechtsspezifische Formulierungen zu
verwenden. Soweit personenbezogene Bezeichnun-
gen nur in männlicher Form angeführt sind, bezie-
hen sie sich auf Männer und Frauen in gleicher Weise.

Geschichte der künstlichen Intelligenz

The History of Artificial Intelligence



Leben im Zeichen des digitalen Wandels

Rasant und unumkehrbar verändern digitale Technologien die Welt, unsere Gesellschaft, unsere Wirtschaft und unser Leben. Vor allem künstliche Intelligenz könnte von größerer Bedeutung sein als die Zähmung des Feuers oder die Elektrizität, meine viele Experten. Die Konsequenzen werden kaum einen Sektor unberührt lassen. Mit der Entwicklung sind zahlreiche Probleme und Gefahren verbunden – aber auch große Chancen, um die Probleme der Menschheit zu lösen. Aufhalten lässt sich die Entwicklung nicht, wir müssen uns ihr stellen und die Rahmenbedingungen proaktiv und zukunftsweisend gestalten. Bildung und Forschung werden darüber entscheiden, was wir aus dem digitalen Wandel machen.

Mit Beiträgen von Hermann Hauser, Helmut Leopold, Stefan Thurner und Helga Nowotny, Manfred Tscheligi, Allan Hanbury, Gerald Reischl u. v. a. wirft das Jahrbuch der Alpbacher Technologiegespräche 2018 einen Blick auf den Status quo und die Zukunftsperspektiven der künstlichen Intelligenz – von konkreten Anwendungsgebieten bis hin zu globalen Fragestellungen in Wirtschaft und Gesellschaft.

Living in times of digital change

Digital technologies are changing the world, our society, our economy, and our life—at great speed and irreversibly. Many experts think that particularly artificial intelligence might have greater significance than the taming of fire or the utilization of electricity. The consequences will leave almost no sector unaffected. The development holds numerous problems and dangers but also great chances to solve the problems of humankind. There is no stopping this development; we must face up to it and proactively create viable conditions to lead us into the future. Education and research will be decisive for what we make of the digital transformation.

Assembling contributions by Hermann Hauser, Helmut Leopold, Stefan Thurner and Helga Nowotny, Manfred Tscheligi, Allan Hanbury, Gerald Reischl, and many others, the yearbook published on the occasion of the Alpbach Technology Symposium 2018 will offer a survey of the status quo and the future perspectives of artificial intelligence—from concrete areas of application to global economic and social issues.



9 783903 207240 >