

DIE BEFREIUNG DES GEFESSELTEN RIESEN

«Wir leben in Homo-Robo-Symbiose»

INTERVIEW RONALD SCHENKEL

Der Physiker und Philosoph Eduard Kaeser gehört zu den wichtigsten Stimmen in der Schweiz, wenn es um unser Verhältnis zur künstlichen Intelligenz (KI) geht. Unsere fortschreitende Angleichung an die Maschinen beschreibt er als fundamentales Problem. Die vor dem Hintergrund gestellte Bildungsaufgabe laute jedoch: «Lernen, wer wir sind.»

Künstliche Intelligenz begleitet uns heute schon im Alltag und doch sprechen wir immer davon, als sei sie Zukunftsmusik. Woran liegt das?

Dominante Technologien brauchen immer die Zukunftsperspektive. Und wenn sie sich auf dem Markt durchsetzen wollen, müssen sie mit entsprechenden Promotionsmitteln gepimpt werden. Dieser Technikglaube kann sich gelegentlich – vor allem in den USA – sogar ins Religiöse steigern. Die technische Zukunft als Erlöserin all unserer Probleme und Sorgen.

Wie intelligent ist KI heute eigentlich wirklich?

In den Anfängen der KI räsionierten die Computerwissenschaftler über die Frage, ob Maschinen, die intelligente Aufgaben übernehmen und bewältigen, selbst auch intelligent sind. Der Clou, oder vielleicht eher die Crux, ist natürlich der Begriff der Intelligenz

selbst. Man kann ihn immer so definieren, dass Maschinen intelligent sind.

Im Besonderen beschäftigte die KI-Forscher die Frage, ob etwas vom Menschen Geschaffenes nicht nur den menschlichen Fähigkeiten ebenbürtig sein, sondern das Menschenmögliche sogar unter Umständen übersteigen könnte. Das ist eine alte, tief philosophische Frage, und fast alle Pioniere der KI bekundeten ein Flair für sie. Dieses philosophische Klima weicht heute einem ökonomischen. Ein Grossteil der jungen Softwaredesigner hält sich kaum mit der Frage auf, ob ihre Kreationen intelligent sind oder nicht, Hauptsache, sie können

sie möglichst schnell auf dem entfesselten Markt der smarten Dinge verhökern.

Aber KI lernt doch inzwischen auch?

Dieser Lernvorgang ist begrenzt und spezifisch. Deep Learning ist wesentlich Statistik. Die statistischen Methoden sind heute äusserst elaboriert, man sollte von ihnen jedoch nicht erwarten, es auf die Stufe der Intelligenz zu schaffen, wie wir sie vom Menschen her kennen. Diese Intelligenz setzt planvolle Intervention und Imagination voraus. Ob wir sie Maschinen vermitteln können, ist eine offene Frage.

KI lernt aufgrund von Datensammlung und Vergleich. Ein Computer braucht 10 000 Vogelbilder, um schliesslich einen Vogel auf Anhieb zu erkennen. Ein Kind kann das viel schneller. Was sagt uns das über die unterschiedlichen Intelligenzen?

«Wir wissen mittlerweile ziemlich gut Bescheid über Rechner, aber wir wissen noch lange nicht, wie das Gehirn funktioniert, und selbst wenn wir es wüssten, bedeutete das nicht, dass das Gehirn ein Rechner ist.»

Primär dies, dass Intelligenz wesentlich zu tun hat mit ganz bestimmten Lernsituationen. Die Intelligenz von Lebewesen ist das Produkt einer langen Geschichte von evolutionären Prüfungen und Herausforderungen. Und diese Herausforderungen sind ja ungeheuer vielgestaltig. Dass sie auch Datensammeln und Vergleichen voraussetzen, ist unbestritten. Aber das ist wahrscheinlich nur ein kleiner Teil der Aktivitäten, dank deren ein Organismus mit seiner Umwelt zurande kommt. Hier muss man noch viel von der Biologie lernen.

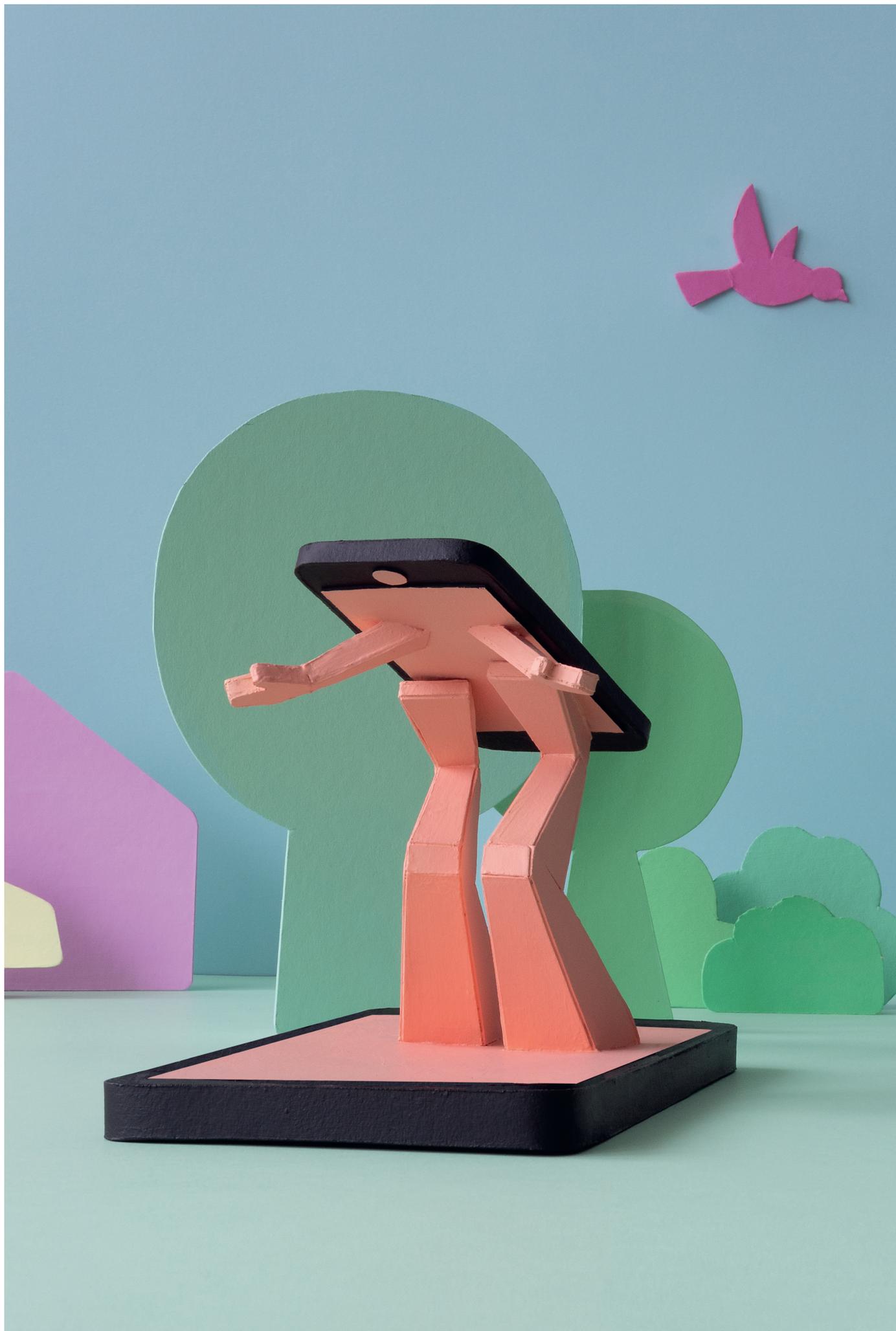
Doch wir gehen davon aus, dass unsere Intelligenz Vorbild für KI sei. Ist das überhaupt richtig? Das ist eine Frage des Paradigmas, der forschungsleitenden Idee. Ein führender KI-Forscher, Stuart Russell, sagt es rundheraus: «Bis jetzt begreifen wir sehr schlecht, warum Deep Learning

so funktioniert, wie es funktioniert. Möglicherweise lautet die beste Erklärung: Weil tiefe Netzwerke tief sind; weil sie zahlreiche Schichten haben, und jede dieser Schichten lernen kann, nach einfachen Regeln einen Input in einen Output zu verwandeln; so dass sich diese simplen Transformationen zu jener komplexen Transformation aufaddieren, die vonnöten ist, um von einer Fotografie zu einer Kategorie zu gelangen.» Die Erklärung wirft einen nicht gerade um. Sie erinnert an die von Molière verspottete «einschläfernde Qualität», die erklären soll, warum Schlafmittel wirken. Eine Tautologie.

Wissen wir eigentlich genug über unser Lernen und unsere Intelligenz, um eine KI zu schaffen, die uns ebenbürtig ist? Sie sprechen damit genau das Problem an. Seit Alan Turing ist die KI-Forschung mesmerisiert von der Aussicht, alles, was der

Mensch tun kann, an Rechenprozesse zu delegieren. Die Deep-Learning-Forscher weisen auch gerne darauf hin, dass das Schichtenmodell neuronaler Netze eine primitive Simulation der Aktivitäten in den Hirnschichten darstelle. Wir wissen mittlerweile ziemlich gut Bescheid über Rechner, aber wir wissen noch lange nicht, wie das Gehirn funktioniert, und selbst wenn wir es wüssten, bedeutete das nicht, dass das Gehirn ein Rechner ist. Der Rechner ist nüchtern gesehen eine heuristische Metapher. Und wenn wir die Metapher mit dem verwechseln, wofür sie steht, dann erliegen wir einer epochalen Missdeutung.

Könnte es dennoch zu einem erfolgreichen Turing-Test kommen? Bestanden ist er ja, wenn Mensch und Maschine ununterscheidbar werden. Turing-Tests werden bisher in einem ganz bestimmten beschränkten Setting durchgeführt. Und darin können KI-Systeme durchaus exzellieren. Die wirklich wichtige Frage ist jedoch, ob KI-Systeme auch in «normalen» alltäglichen Umwelten bestehen können, also in Settings, die mit Ambiguitäten, unscharfen – «fuzzy» – Bedingungen, unerwarteten Eventualitäten gespickt sind. Die KI-Forscher sprechen in diesem Zusammenhang vom «Rahmenproblem» (frame problem). Unsere Handlungen finden immer statt in einem Rahmen voller impliziter, unbewusster Annahmen, auf die wir in «fuzzy» Situationen zurückgreifen können. Dieser flexible, nie verfestigte



Rahmen bildet das, was man die praktische oder die Commonsense-Intelligenz nennt. Wie implementiert man so etwas in KI-Systemen? Wie können die Commonsense haben? Das ist der entscheidende Test. Im Übrigen irrte Turing, was das Lernen anbelangt. Er schrieb, dass das Kinderhirn so etwas sei wie ein leeres Heft, das man beim Schreibwarenhändler kauft: «Ziemlich wenig Mechanismus und viele leere Blätter.» Nichts könnte falscher sein.

Oder ist es nicht wahrscheinlicher, dass KI sich zu einer völlig eigenständigen Form von Intelligenz entwickelt, die wir am Ende gar nicht mehr verstehen? Tatsächlich befinden wir uns bereits in einer Entwicklungsphase, die ich das Zeitalter der unbegreiflichen Maschine nenne. Neuronale Netzwerke bestehen oft aus Millionen von Elementen, die einen numerischen Input in einen numerischen Output verwandeln. Das heisst, das System entwickelt womöglich Regeln, die nur es selbst kennt. Der Designer hat be-

grenzten – wenn überhaupt – Einblick in das, was sich im Innern abspielt. Mit zunehmender Schichttiefe wird das KI-System selbständiger: eine Black Box. Sie funktioniert oft überraschend gut und liefert Voraussagen mit erstaunlicher Genauigkeit – um den Preis, dass man nicht versteht, wie sie zu diesem Resultat gekommen ist. Es manifestiert sich ein umgekehrtes Verhältnis von genauer Voraussage und Verstehbarkeit: Je fähiger das System zu exakter Voraussage ist, desto schwieriger ist es interpretierbar. So gesehen, haben wir die Wahl zwischen durchschaubaren, aber ineffizienten und undurchschaubaren, aber effizienten Orakeln.

Kommen wir an den Anfang zurück: KI ist bereits mit und unter uns. Und sie hat unseren Alltag und auch unser Verhalten verändert.

Wie wird das weitergehen? Das weiss niemand genau. Wir sollten jedenfalls auf die nüchternen KI-Leute hören. Gemäss Michael Negnevitsky etwa, der ein autoritatives Lehrbuch über KI geschrieben hat, befinden wir uns bei lernenden Maschinen im

Stadium des Papierflugzeugs im Vergleich zum Überschallflieger. Und ohnehin: Wer aus dem Siegeszug der smarten Geräte schliesst, in absehbarer Zeit würde sich eine von uns unabhängige Superintelligenz entwickeln, erinnert an den Affen, der auf den Baum steigt und prahlt, den ersten Schritt zum Mond getan zu haben.

Und wird das weiterhin so schleichend passieren, dass wir gar nicht merken, was vorgeht? Die digitalen Technologien sind im Wesentlichen Technologien des alltäglichen Verhaltens. Die Verhaltensökonomik hat als Disziplin Hochkonjunktur. Sogar einen Nobelpreis hat sie schon erhalten (Daniel Kahnemann, Amos Tversky). Es geht um ein soziokulturelles Problem: unsere Angleichung an die Maschinen. Wir haben zwar bis heute keine Antwort auf die Frage, ob Maschinen denken können, aber wir gewöhnen uns allmählich an die Sprache der Computeringenieure und Software-designer, die ja nicht selten von den Funktionen der Computer so sprechen, als würden sie denken. Die Metaphorik hat sich sozusagen in die Umgangssprache eingeschlichen, und dadurch wird der Eindruck erweckt, der Computer sei buchstäblich zur Intelligenz erwacht. Die Populärliteratur ist voll von solchen undurchdachten Metaphern. Sie spricht andauernd von den Computern, die immer mehr menschliche Tasks übernehmen, so wie umgekehrt der menschliche Körper als «biologische Maschine» neuerdings als «biologischer Algorithmus» betrachtet wird. Das heisst, die Maschine beginnt, sich sozusagen in unseren technisierten Lebensformen einzubürgern als ein Quasi-Subjekt. Wir brauchen sie nicht bloss, wir leben mit ihr zusammen wie mit Mitmensch und Tier. Wir leben in Homo-Robo-Symbiose. Und das Risiko

LERNEN UND KI

Acht angehende Illustratorinnen und Illustratoren, Studierende der Schule für Kunst und Design Zürich, haben sich im Rahmen eines Projekts mit der Thematik «Lernen und Künstliche Intelligenz» auseinandergesetzt. Die von den beiden Dozentinnen Manuela Monn und Martina Walthert gestellte Aufgabe verlangte eine Visualisierung der in diesem Interview angesprochenen Probleme. Mit ihrer Arbeit «Flugi-Spiel» hat Carmela Leggio diese Aufgabe am besten gelöst.

www.carmela-leggio.ch

ist durchaus da, dass wir den Roboter nicht anthropomorph betrachten, sondern uns Menschen robomorph.

Müssen wir uns Sorgen machen?

Sorgen machen müssen wir uns weniger über die KI-Systeme als vielmehr über uns selbst. Wir adaptieren uns nämlich viel leichter an die neue Technologie, als sich diese an uns adaptiert. Wir sind ja schon heute Mischwesen – Hybride – aus Mensch und Gerät. Man schaue nur auf die alltäglichen Gepflogenheiten des kommunikativen Umgangs. In der kürzesten Zeit ist dieses handliche Gerät, das die meisten mit sich herumtragen, sozusagen in unsere Psyche eingesunken. Es bestimmt unsere sozialen Umgangsformen, das Selbstverhältnis und -verständnis. Nebenbei bemerkt: Die Science Fiction ist fasziniert von extraterrestrischen Wesen, die ein uns überlegenes Intelligenzniveau aufweisen. Dabei sind die Aliens bereits unter uns, in der Gestalt hochentwickelter automatischer Systeme. Ein Grossteil des globalen Finanzmarktes wird von solchen «Aliens» gesteuert, in deren unerforschliche Ratschlüsse man kaum noch Einblick hat. Es gibt Schachprogramme mit Strategien, denen selbst Grossmeister schwer, wenn überhaupt auf die Schliche kommen. Sie muten wie eine «torkelnde» Intelligenz an, die scheinbar unlogisch, in zufallsgesteuerter Weise ihre Schritte vollführt

EDUARD KAESER wurde 1948 in Bern geboren. Er studierte theoretische Physik, anschliessend Wissenschaftsgeschichte und Philosophie an der Universität Bern. Bis 2012 war er Gymnasiallehrer für Physik und Mathematik. Er publiziert über Themen zwischen Wissenschaft und Philosophie. Neben Büchern veröffentlicht er Essays und Artikel unter anderem in der «Wochenzeitung», der «Zeit» und der «Neuen Zürcher Zeitung».

«Sorgen machen müssen wir uns weniger über die KI-Systeme als vielmehr über uns selbst.»

und erst noch regelmässig den Gegner schachmatt setzt. Fremd und unbegreiflich, wie ein aussergalaktisches Wesen.

Was sollten wir normalsterblichen Nutzer lernen, um mit KI richtig umgehen zu können?

Es gibt darauf eine simple Antwort: Lernen, wer wir sind. Aber sie ist nur auf Anhieb simpel. Das avancierte Verständnis der Maschine wirkt auf unser Selbstverständnis zurück als die Frage: Was die Maschine auch kann, worin besteht der Unterschied zwischen menschlicher und maschineller Ausführung eines Könnens? Das ist, meine ich, die zentrale Frage zu Beginn des 21. Jahrhunderts – das Vermächtnis Alan Turings. Und sie wird nicht primär von den Computingingenieuren beantwortet werden, sondern von

den Anthropologen, das heisst: letztlich von jedem einzelnen von uns. Hoffen wir das zumindest.

Glauben Sie, diese Bildungsaufgabe wird erkannt und umgesetzt?

Es gibt durchaus Anzeichen, dass in den Schulen – etwa im Gymnasialunterricht – die Thematik erkannt wird. So habe ich schon Informatik- und Philosophielehrer zusammenarbeiten sehen. Wichtig sind in diesem Kontext die Fachhochschulen, in denen nicht einfach tunnelisiert technisches Know-how vermittelt werden sollte. Das Interesse der Fachhochschulen ist gross.

Wie wird umgekehrt KI unser Lernen verändern? Werden Hologramme bald die neuen Trainer sein?

Nein. Lehrer und Trainer werden ihre Funktion in der Konstellation der digitalen Intelligenz neu definieren müssen. Das gehört ja auch zur Bildungsaufgabe. Wie gesagt: Was wissen wir Menschen eigentlich über unsere Kompetenzen?

Was wird KI letztlich aus uns machen?

Kann ich nicht beantworten. Es gibt beide Wege. Wie einer der klügsten und einflussreichsten Kritiker der KI, der amerikanische Philosoph Hubert Dreyfus, schon vor einem halben Jahrhundert warnte, liegt die Gefahr nicht so sehr in superintelligenten Maschinen als vielmehr in subintelligenten Menschen. Die Warnung ist aktueller denn je.