

Liebe Maschinen – lasst uns doch Freunde sein

Bei der Weiterentwicklung der künstlichen Intelligenz ist jetzt psychologisches Fachwissen gefragt. Die Maschinen benötigen eine Beziehungstherapie.

Stefan Betschon, NZZ 12.2.2019



Hätten die Astronauten nur auf die Maschinen gehört, wären die Mondlandungen – im Bild die Mission Apollo 15 – möglicherweise nie geglückt. (Bild: Nasa/Corbis/Getty)

Kurz vor der Landung, zehn Kilometer über der Mondoberfläche, begann neben dem Computerbildschirm ein Lämpchen zu blinken. «Wir haben ein Computerproblem», funkte der Pilot der Apollo-Landefähre, Neil Armstrong, zur Erde. «Fehlercode 1202.» Im Kontrollzentrum in Houston versuchten Computerfachleute unter grösstem Zeitdruck herauszufinden, was es mit diesem Fehlercode auf sich hatte. Schliesslich sagte einer: «Wir gehen weiter mit diesem Alarm.» Für diesen Funkspruch, für den Entscheid, den Computer zu ignorieren und den Landeanflug fortzusetzen, wurde der Ingenieur im Kontrollzentrum später vom amerikanischen Präsidenten mit einem Preis ausgezeichnet. Wegen Gesteinsbrocken auf dem geplanten Landeplatz zu einem

Umweg gezwungen, Fehlermeldungen bezüglich Treibstoffmangels ignorierend, brachte Armstrong die Fähre schliesslich auf den Boden und konnte die historischen Worte sprechen: «The Eagle has landed.»

Pilot gegen Autopilot

Tausende von Ingenieuren haben in den USA während der 1960er Jahre für die Raumfahrt gearbeitet, sie haben in vielen Bereichen wichtige Innovationen vorangetrieben. Auch die Kooperation von Pilot und Autopilot musste grundlegend überdacht werden. Es gab Techniker, die die Astronauten als eine Art Frachtgut betrachteten; sie glaubten, dass der Erfolg der Apollo-Missionen am besten garantiert werden könne, wenn die «humans out of the loop» blieben, wenn die Menschen an Bord während des Fluges von allen wichtigen Entscheidungen ausgeschlossen würden. Andere vertraten den Standpunkt, dass eine bemannte Raumfahrt ohne Männer, dass eine Entdeckungsreise im All ohne Entdecker nicht sinnvoll sei. Vor allem die Piloten wollten sich nicht zu Passagieren degradieren lassen: Sie waren alle hochdekorierte Militärpiloten, die Kampfeinsätze geflogen hatten. Die Ereignisse bei der ersten Mondlandung gaben ihnen recht: Es braucht die «humans in the loop», Mensch und Maschine sind aufeinander angewiesen.

Das Apollo-Programm hat der Informatik – von der Halbleitertechnik bis zum Software-Engineering – viele wichtige Impulse gegeben. Doch eine zentrale Lektion scheint vergessen gegangen zu sein: Eine künstliche Intelligenz, die nicht in der Lage ist, sich mit natürlicher Intelligenz zu verbünden, ist zum Scheitern verurteilt.

Viele spektakuläre Flugunfälle der jüngeren Vergangenheit – beispielsweise 2009 der Absturz eines Airbus A330 der Air France über dem Atlantik mit 216 Toten – sind die Folge von Unstimmigkeiten zwischen Pilot und Autopilot. Wo die Menschen vom Autopiloten zur Untätigkeit verurteilt werden, ist es unvermeidlich, dass ihre Fähigkeiten verkümmern. Wenn dann der Autopilot versagt, kann es geschehen, dass die Piloten von einer Situation, die jeder Anfänger meistern könnte, überrascht und überfordert werden.

Spektakuläre Unfälle

Autonom fahrende Autos waren in jüngster Vergangenheit in einige spektakuläre Unfälle verwickelt, bei denen Menschen starben. 2016 wurde in Florida ein Tesla-Fahrer enthauptet, weil der Autopilot einen weissen Sattelschlepper, der die Fahrbahn versperrte, nicht als Hindernis erkannte. Nach einer Untersuchung nannte die amerikanische National Traffic Safety Administration menschliches Versagen als Unfallursache. Der Autopilot setze die ständige Aufmerksamkeit des Fahrers voraus, doch dieser habe sich während der Fahrt einen Harry-Potter-Film angeschaut. Das

Problem mit den Autopiloten ist nicht nur, dass sie manchmal Fehler machen, das Problem ist auch, dass sie die Menschen dazu verleiten, Fehler zu machen.

Elon Musk hatte 2016 versprochen, dass ein Tesla 2017 ohne Chauffeur von Los Angeles nach New York fahren werde. 2017 bringe für das autonome Fahren den Durchbruch, erklärte auch John Krafcik, der innerhalb des Alphabet-Konzerns die Firma Waymo leitet. Der Durchbruch lässt allerdings auf sich warten. Die Technik ist noch nicht bereit, die Software muss noch dazulernen. Sie muss nicht nur die Strassen genauer studieren, sie braucht Nachhilfe auch im Fach «Menschliches und Allzumenschliches». Denn selbst wenn es in einem einzelnen Auto gelänge, den Menschen aus dem «Loop» zu nehmen und ruhigzustellen, müsste sich die Software doch auf die Begegnung mit Menschen einstellen: Kindern, die auf die Strasse rennen, Velofahrern, die sich zwischen den Autos hindurchschlängeln, Autofahrern, die vergessen, dass sie Vortritt haben.

Oft heisst es, die Computer seien «nachweislich» die besseren Autofahrer. Doch wo wurde dieser Nachweis geführt? Wie könnte man diesen Nachweis überhaupt erbringen? Wie viele Kilometer muss eine Software unfallfrei bewältigt haben, bis sie als sicher gelten darf? Software, die im Rahmen von Machine-Learning-Prozessen hergestellt wurde, zeigt sich dem Menschen als Blackbox. Diese Software besteht nicht aus Codezeilen, die man lesen, nicht aus Algorithmen, deren Zuverlässigkeit man mathematisch beweisen könnte. Manchmal überrascht einen diese Software durch ihre Intelligenz, manchmal durch ihre Dummheit. Amerikanische Forscher haben 2017 gezeigt, dass einige wenige Zentimeter weisses Klebband genügen, um die Software eines autonom fahrenden Autos zu verwirren. Auf ein Stoppschild geklebt, können ein paar wenige weisse Punkte und Striche, die ein Mensch leicht übersieht, die Software dazu bringen, Gas zu geben, anstatt abzubremesen.

Neue Arbeitswelten

Fortschritte bei der Erforschung der künstlichen Intelligenz (KI) werden oft als Schaukampf zwischen Mensch und Maschine inszeniert. 2016 stellte die Google-Tochter Deepmind eine Alphago genannte Software ins Rampenlicht der Weltöffentlichkeit: In einem mehrtägigen Event mit mehreren Begegnungen konnte das künstliche neuronale Netzwerk den besten Go-Spieler der Welt schlagen. Viel Beachtung fand 1997 der Kampf zwischen dem Computer Deep Blue von IBM und Garri Kasparow. Der Schachweltmeister verlor, die Maschine triumphierte.

Die Berichterstattung über KI folgt dem Muster der Schaukämpfe, es geht um ein «wir gegen die». Alle grossen Beratungsunternehmen haben sich schon zu Wort gemeldet, Studie um Studie versucht darzulegen, welche Berufsleute bald vom Spielbrett weggespickt werden. Laut einer vielzitierten Studie der University of Oxford ist in den USA knapp die Hälfte aller Jobs bedroht. Andere Studien geben sich mit weniger

zufrieden. Der Prozentsatz der gefährdeten Arbeitsstellen mag variieren, die Resultate unterscheiden sich, doch bei der Herangehensweise ähneln sich all diese Studien: Es geht um ein «wir gegen die».

Paradigmenwechsel

Software-Ergonomie und Interface-Design sind Forschungsthemen, die innerhalb der Computerwissenschaft eine lange Tradition haben. Als Ahnherr dieser Forschungstradition gilt der amerikanische Elektroingenieur Doug Engelbart. Er hat unter anderem die Computermouse erfunden. Er arbeitete in den 1960er Jahren am Stanford Research Institute im Silicon Valley nur wenige hundert Meter entfernt vom Artificial Intelligence Laboratory der Stanford University. Hier hatte John McCarthy das Sagen. Er gilt als Gründungsvater der KI. Es heisst, dass Engelbart und McCarthy sich nicht ausstehen konnten. Als bei einer Begegnung McCarthy von seiner Mission berichtet habe, die Maschinen intelligent zu machen, habe Engelbart geantwortet: «Was? Das alles machen Sie für die Maschinen? Und was machen Sie für die Menschen?»

Engelbart wollte die Computertechnik dazu nutzen, die Fähigkeiten des Menschen zu erweitern. Sein Labor nannte er Research Center for Augmenting Human Intellect. Die Intelligenzerweiterung (Intelligence Augmentation, IA) gilt als Gegenentwurf zur Artificial Intelligence (AI). Er habe die Wissenschaft der KI begründet im Bemühen, von der Psychologie wegzukommen und die Kybernetik zu überwinden, schrieb McCarthy einmal. Während Jahrzehnten dominierte der von ihm propagierte symbolische Ansatz weltweit die Erforschung der KI. Doch auch als dieses Paradigma in den 1990er Jahren verdrängt werden konnte, blieb der KI die Menschenferne erhalten.

Es ist absehbar, dass die KI die Welt und auch die Arbeitswelt stark verändern wird. Doch dieser Umbruch wird sich nicht anhand von Arbeitslosenstatistiken erfassen lassen. Die künstlich intelligenten Maschinen werden sich im Berufsleben vieler Menschen bemerkbar machen, aber nicht als Konkurrenten, sondern als Kollegen. Um sich bei der «kollaborativen Intelligenz» zu bewähren, müssen die Maschinen anders werden. Sie müssen lernen, das menschliche Gegenüber bei der Lösungsfindung einzubeziehen, sie müssen lernen, sich selber zu hinterfragen, mit Unsicherheiten umzugehen, zu improvisieren. Sie werden nicht nur Lösungen liefern müssen, sondern auch Lösungswege, nicht nur Antworten, sondern auch Fragen.