



„Vielleicht haben Maschinen eines Tages Bewusstsein“

Maschinen übernehmen immer mehr Aufgaben von Menschen. Sie bestimmen zunehmend unser Verhalten und unsere Entscheidungen – bis hin zur Entscheidung über Leben und Tod. absatzwirtschaft sprach mit dem Schweizer Philosophen und Maschinenethiker Prof. Oliver Bendel über die Algorithmisierung des Marketings und Maschinen als Träger moralischer Entscheidungen

Von
Peter Hanser

Herr Prof. Bendel, über welche Roboter reden wir im Zusammenhang mit Roboter- und Maschinenethik? Den niedlichen Roboter, der Gäste an der Hotelrezeption empfängt, nichthumanoide Roboter, autonom fahrende Autos, Börsenprogramme oder Selbst-Check-out-Kassen im Handel?

OLIVER BENDEL: Am Anfang, als der Begriff „Roboter“ aufkam, hat man damit vor allem humanoide Maschinen gemeint. Das hat sich wieder verflüchtigt. Den Begriff des Roboters fasse ich sehr weit. Trotzdem würde ich mich dagegen wehren, jede Maschine gleich Roboter zu nennen. Wir können von spezifischen Maschinen sprechen, die sich bewegen oder sich verhalten können, oder von autonomen und teilautonomen Maschinen, wenn es um Maschinenethik geht. Industrieroboter oder Serviceroboter wie Saug- und Mähroboter sind Roboter. Neben solchen Bewegungsautomaten würde ich auch Bots und Agenten als Roboter bezeichnen, als Softwareroboter. Aber bestimmte Maschinen wie Kaffeefullautomaten sind keine Roboter.

Das Roboy-Projekt: Das Projekt des Artificial-Intelligence-Labors der Universität Zürich hat sich das ambitionierte Ziel gesetzt, einen Roboter zu bauen, der sich genauso bewegt wie ein Mensch. Statt Motoren in den Gelenken hat Roboy Junior – der erste Prototyp eines solchen Roboters – Muskeln und Sehnen

„Wir sollten nicht Maschinen wollen, die über Leben und Tod von Menschen entscheiden“

Oliver Bendel



Roboter sind auf dem Vormarsch. Welches sind die Treiber dieser Entwicklung? Technikbegeisterung, Rationalisierung oder Gewinnmaximierung?

Alles zusammen. Was die Industrierobotik angeht, natürlich Gewinnmaximierung. Die große Idee der Industrie 4.0 war, die ins Ausland abgewanderte Produktion wieder ins Land zu holen oder diese hier zu halten. Zudem gibt es alle möglichen Komponenten wie Sensoren oder Kameras, die sehr günstig geworden sind. Das bedeutet, dass man einfache Roboter für sehr wenig Geld zusammenbauen kann. Ein weiterer Treiber ist der 3-D-Druck. Zum Beispiel sind bei Roboy, einem Produkt von Rolf Pfeifer und anderen, die Knochen, an denen die künstlichen Muskeln haften, aus dem 3-D-Drucker. Und es gibt den „Silicon-Valley-Treiber“, der bestimmte Technologien pushen und uns von diesen abhängig machen will. Es existieren also handfeste Interessen der deutschen und der europäischen Wirtschaft. Außerdem darf die Verspieltheit der Entwickler nicht außer Acht gelassen werden.

Steigt damit die Notwendigkeit, sich mit dem Thema Ethik auseinanderzusetzen?

Aus verschiedenen Richtungen, ja. Ich unterscheide Maschinenethik als Gestaltungsdisziplin und andere Disziplinen wie Informationsethik oder Technikethik als Reflexionsdisziplinen. In der Maschinenethik wollen wir wirklich Maschinen konzipieren und am Ende prototypisch bauen. Dabei arbeiten wir eng mit KI, Robotik und anderen Disziplinen zusammen. In der Informationsethik und der Technikethik reflektieren wir Probleme, die sich beim Einsatz von Robotern ergeben. Es ist beides sinnvoll und notwendig. Wir haben zum ersten Mal in der Geschichte der Ethik eine Form der Ethik, die danach fragt, wie man maschinelle Moral umsetzen kann. Zugleich haben wir zunehmende Probleme durch den Einsatz von Robotern, etwa durch den Ersatz von Arbeitskräften, Kollisionen von Menschen und Maschinen und technikbezogene Sexpraktiken. Es stellt sich die Frage, was wir künftig mit Maschinen tun wollen. Damit sind dann Bereichsethiken

gefordert wie Informations-, Technik-, Medizin-, Sexualethik und so weiter.

Geht es auch um die Fremdbestimmung durch die Technik?

Richtig. Das ist ein großes Thema der Technikethik und auch der Informationsethik. Wie abhängig sind wir von diesen Technologien, was ist mit unserer informationellen Selbstbestimmung, wohin gehen unsere Daten, wie stark beobachten die Maschinen uns? Das führt zu Fragen wie „Müssen die Maschinen alle vernetzt sein?“ oder „Müssen sie alle mit Servern verbunden sein?“. Nicht alle Roboter müssen Datensauger sein, aber viele sind es, wie die Puppe Hello Barbie von Mattel – das ist ein Datensauger par excellence – oder Echo von Amazon. Es existieren allein im auditiven Bereich inzwischen fünf, sechs, sieben Geräte in jedem Haushalt, die uns belauschen können. Bei den Browsern können wir weitermachen. Von Google wurde jetzt bekannt, dass es bei Chrome den Suchverlauf für personalisierte Werbung nutzen will. Von daher gibt es alle möglichen Fragen

aus der Ethik heraus. Die einen betreffen eher die Gestaltung, die anderen eher die Nutzung der Maschinen.

Was beinhaltet der Begriff „Roboterethik“?

Maschinenethik habe ich eingeführt als Pendant zur Menschenethik. Was beide unterscheidet, ist das Subjekt der Moral. Bei der Menschenethik ist das Subjekt der Moral der Mensch, bei der Maschinenethik die Maschine. Dann können wir noch nach dem Objekt fragen. Das Objekt kann der Mensch oder das Tier sein. In der Maschinenethik fragen wir nach der Möglichkeit moralischer Maschinen. Wir können das metaphorisch machen oder zielhaft, nach dem Motto: Unser Ziel in 100 Jahren ist die moralische Maschine. Bei der KI läuft die gleiche Diskussion seit 50 Jahren: Dürfen wir den Begriff der Intelligenz auf Maschinen anwenden? Ich würde ja sagen, wir sind frei, das zu tun. Wir können uneigentlich sprechen oder uns ein Ziel vorgeben, nämlich intelligente Maschinen zu schaffen. Dann habe ich mit solchen Wörtern wenig Probleme. „Roboterethik“ ist dagegen ein schillernder Begriff. Es gibt drei Richtungen: Entweder man betreibt Roboterethik als Teil der Maschinenethik. Dann lässt man bestimmte Maschinen weg, die keine Roboter sind, zum Beispiel 3-D-Drucker. Maschinenethik kann sich aber auch auf 3-D-Drucker beziehen, wenn man normative Ansätze für diese verwendet, die ihnen beispielsweise verbieten, Waffen auszudrucken. Man sieht, nicht alles, was Maschinenethik anfasst, hat mit Robotern zu tun. Entweder man begreift Roboterethik als Teil der Maschinenethik – oder als Bereichsethik, die sich mit den Chancen und Risiken des Einsatzes von Robotern befasst. Die dritte Bedeutung resultiert aus der Frage, ob Roboter Rechte haben.

Warum sollten Roboter in das moralische Universum aufgenommen werden?

Wegen ihrer Teilautonomie oder Autonomie. Es sind Maschinen, die sich selbstständig bewegen, von uns nicht direkt gesteuert werden und allein gelassen bestimmte Entscheidungen treffen müssen, auch moralischer Art. Die Entscheidungen, die sie treffen müssen, können sich auf Situationen beziehen, die wir voraussehen, oder die Maschine lernt dazu, ist selbstlernend. Dann kann sie sich nach und nach ein bestimmtes Wissen aneignen, Regeln bilden und entscheiden. Das ist potenziell gefährlich. Die Maschine kann Subjekt der Moral sein, aber es ist völlig klar, es ist ein ganz anderes moralisches Subjekt als der Mensch. Die Maschine hat kein Bewusstsein – wir können sehr vorsichtig von Absichten sprechen –, keine Gefühle, keine Empathie.

Können Roboter Träger moralischer Rechte sein? Vermenschlichen wir damit nicht den Roboter?

Was die Roboter als Subjekte angeht, habe ich wenig Mühe. Sie können beispielsweise Träger von Pflichten sein. Ich halte sie aber nicht für Träger von Rechten. Damit kommen wir zu Maschinen als Objekten. Für mich sind Träger von Rechten empfindungsfähige, leidensfähige oder lebenswillige Entitäten. Roboter gehören nicht dazu. Sie haben keine Rechte und können in den nächsten 30 bis 40 Jahren keine Rechte erhalten. Warum bin ich so zögerlich? Ich glaube, dass Roboter bis auf Weiteres, wenn sie nur maschinell gestaltet werden,

keine moralischen Objekte sein können. Wir können sie behandeln, wie wir wollen: Wir können sie versklaven, treten und was immer wir wollen.

Und wenn sie entscheiden, was sie tun können?

Dann betrifft das wiederum die Maschinen als Subjekte. Sie werden damit nicht zu moralischen Objekten. Selbst den Roboter, der moralisch gut handelt, dürfen wir treten, steinigen, schlagen. Es sei denn, wir bemühen noch ein anderes Argument, nämlich das der Verrohung. Dieses hat Konjunktur seit der Antike, und auch Kant hat es benutzt. Wenn wir Tiere schlagen, so die Denkweise, ist dies wegen der Tiere egal, aber wir verrohen, wir behandeln andere Menschen oder uns selbst auf unbotmäßige Weise. Auch wenn wir Maschinen schlagen, könnten wir verrohen. Was die Objekte der Moral angeht, muss ich nochmals weiter ausholen. Es gibt zwei Arten von Cyborgs. Der übliche Cyborg ist beispielsweise ein Mensch mit einer Hightech-Prothese oder einem Herzschrittmacher. Es gibt noch eine andere Form: Man nimmt technische Strukturen – das ist der umgekehrte Weg – und versieht sie mit Nervenzellen oder Muskeln. Wenn wir da weitermachen – was ich für heikle Forschung halte –, könnten wir eines Tages durch das Einpflanzen von Ganglienknoten oder kleinen Gehirnen bei bewussten Maschinen landen. Eine solche Forschung könnte auch zur Leidensfähigkeit oder Empfindungsfähigkeit von Maschinen führen.

Kann man programmieren, dass der Roboter in diversen Situationen vernünftige Entscheidungen trifft?

Das kann man programmieren, indem man sehr viele Situationen antizipiert, in die sich die Maschine begeben könnte. Das ist im Moment das große Thema bei autonomen Autos. Bei Tesla hat man eine Situation nicht vorausgesehen, was zu einem tödlichen Unfall führte. Die Automobilindustrie ist zurzeit sehr vorsichtig bei ihren Ansätzen. Sie sagt, selbstlernende Autos seien im Moment zu gefährlich. Es sind Millionen von Situationen zu bedenken und wie wir feststellen müssen, können wir bestimmte Situationen nicht oder nicht präzise genug voraussehen.

Wir können selbstlernende Maschinen bauen, die Situationen kennenlernen, daraus gewisse Schlussfolgerungen ableiten und selber Regeln bilden. Das ist sehr gefährlich, weil sich die Maschine dann von



Vita

Oliver Bendel, geboren 1968 in Ulm, hat an der Universität Konstanz studiert (Philosophie, Literaturwissenschaft sowie Informationswissenschaft) und an der Universität St. Gallen promoviert. Er arbeitete als Projektleiter und leitete technische und wissenschaftliche Einrichtungen, unter anderem das Competence Center E-Learning an der Universität St. Gallen und die Arbeitsgruppe Business Communication am Fraunhofer ISST in Dortmund. Im April 2009 wurde er von der Fachhochschule Nordwestschweiz zum Professor ernannt. Bendel ist Sachverständiger des Deutschen Bundestags und Mitglied beim International Center for Information Ethics (ICIE). Neben seiner Professur ist er als freier Schriftsteller tätig und hat mehrere Gedichtbände und Romane veröffentlicht.



© Daimler

Busfahrt in die Zukunft: In Amsterdam testet Daimler den ersten Bus, der automatisiert fahren kann, ohne dass ein Fahrer eingreifen muss. Der Bus mit Citypilot fährt auf speziell ausgewiesenen Busspuren

dem sehr weit entfernen kann, was die Entwickler eigentlich wollten, wie bei Tay, einem Bot von Microsoft. Der kam in schlechte Gesellschaft und am Ende des Tages war er ein Rassist, bejubelte Hitler und beschimpfte die Juden. Es gibt noch eine andere, eine regel- und fallbasierte Vorgehensweise: Wir geben der Maschine entweder bestimmte Regeln vor, nach denen sie sich verhalten soll, oder wir füttern sie mit Fällen und sie vergleicht die Situation mit Fällen, die sie bereits in ihrer Datenbank hat.

Belegt das Beispiel Tay, dass sich selbstlernende Systeme nicht kontrollieren lassen?

Sie geraten immer mehr außer Kontrolle. Bei Tay hat man versäumt, den Bot zu „deckeln“. Zum Beispiel hätte Microsoft Metaregeln finden können wie „Die Maschine bejubelt niemals Diktatoren“ oder „Adolf Hitler wird niemals positiv bewertet“. Wenn man selbstlernende Maschinen entwickelt, würde ich dazu raten, sie mit Metaregeln zu versehen. Aber prinzipiell sind sie gefährlich.

Können befriedigende Regeln und Formeln in Bezug auf Leben und Tod von menschlichen Verkehrsteilnehmern entwickelt werden?

Ganz schwierig. Ich rede gerne von einfachen moralischen und komplexen moralischen Maschinen. Wo die Sache toll funktioniert, ist in geschlossenen oder halb offenen Umgebungen. In meinem Haushalt kann der Saugroboter bestimmte moralische

Entscheidungen treffen, etwa wenn er Spinnen und Marienkäfer nicht einsaugt.

Aber wenn ich die Tierchen gar nicht im Haus haben will?

Da beginnen die Details. Sie sagen als Verbraucher, diese Maschine will ich oder ich will die Maschine, die einen Killbutton für Spinnen hat, aber Marienkäfer am Leben lässt. In diesem Kontext funktioniert es. Das regelt sogar der Markt. Manche Verbraucher werden die Maschinen selber anpassen über ein bestimmtes Menü, sofern dieses vorgesehen ist. In der offenen Welt wird es schwieriger.

Deshalb befürchte ich, dass Bundesverkehrsminister Alexander Dobrindt mit der Ethikkommission für autonomes Fahren nicht sehr weit kommen wird. Es ist durchaus klug, was er bisher gesagt hat. So möchte er selbstständig fahrende Autos mit zwei Metaregeln deckeln. Erstens: Sachschaden ist immer Menschenleben vorzuziehen. Damit beginnen wir uns schon in Richtung der Maschinenethik zu bewegen. Zweitens: Der Roboter soll nicht qualifizieren: Der Minister ist gleich viel wert wie der einfache Bürger. Darauf kann man sich verständigen. Das ist klug. Beim Rest wird es kompliziert. Die Frage ist, was macht die Maschine, wenn die Bremsen versagen, was macht sie außer qualifizieren? Interessanterweise hat er nicht gesagt, dass sie nicht qualifizieren darf. Ich fürchte, darauf läuft es hinaus.

Das heißt, Personen, zum Beispiel alt oder jung, einen Wert beizumessen?

Die Maschine zählt einfach durch. Sie wird feststellen, da ist eine Person, dort sind fünf Personen, also nehme ich das kleinere Übel. Damit sind wir beim klassischen Trolley-Problem. Die Ethikkommission darf die Maschine nicht qualifizieren lassen, sie muss Sachschaden vor Menschenleben stellen, aber was macht man dann in bestimmten Situationen? Sie wird vielleicht die Maschine, wie gerade beschrieben, quantifizieren lassen. Aber was ist, wenn die eine Person Sie sind und die anderen fünf Personen IS-Kämpfer? Dann ist die quantitativ abgestützte Entscheidung offensichtlich falsch. Wenn wir das auf die Spitze treiben, ergibt es wieder Sinn. Wir dürfen aus moralischer Sicht sogar Passagierflugzeuge abschießen, die im Begriff sind, in ein Atomkraftwerk zu steuern. Wir müssen leider die Entscheidung treffen, weil wir dann 200 Menschen für 20 000 opfern.

Wie sollte ein Gremium aufgestellt sein? Wer definiert, was moralisch ist?

Erst einmal gehören Hersteller und Programmierer hinein. In eine Ethikkommission gehören auch Ethiker – aber da beginne ich schon zu schlucken, wenn ich an den Deutschen Ethikrat denke, der tendenziös aufgestellt ist. Das Letzte, was ich Ihnen antworten würde, wäre, dass Maschinenethiker diese Regeln finden. Sie wollen mehrheitlich nicht konkrete moralische Entscheidungen treffen, sie konzipieren die Möglichkeiten, die Bedingungen, gewährleisten das Funktionieren der moralischen Maschine. Ich vermute, dass eine Mischung



Logistik der Zukunft: Im Förderprojekt Villa Ladenburg der Daimler und Benz Stiftung wird untersucht, inwieweit vollautomatisches Fahren im Güterverkehr auf der Straße sinnvoll und praktisch umsetzbar ist

aus Wirtschafts- und Fachvertretern sowie ein paar Philosophen und leider noch mehr Theologen berufen wird. Sie werden die Probleme aber nicht lösen. Eigentlich müsste die Gesellschaft mitdiskutieren. Diese könnte am Schluss aber kaum befriedigende Formeln finden in Bezug auf Leben und Tod von Menschen. Denn das sind individuelle Akte, bei denen nicht nur die Gesellschaft ein Interesse hat, sondern auch die Mutter oder der Vater des Kindes, das totgefahren wird.

Es gibt ethische Unterschiede zwischen Kulturen. Werden dann noch unterschiedliche moralische Settings in verschiedenen Ländern benötigt?

Sie sprechen mit moralischen Settings einen wichtigen Begriff an. Moral kann man charakterisieren als normatives Setting. Dieses besteht aus Ideen, Werten, Überzeugungen, Vorstellungen, rationalen Überlegungen in Bezug auf das, was gut und böse ist. Die Maschine hat ein abgespecktes Setting, das sich je nach Anwendungsgebiet, je nach Kultur und so weiter unterscheidet. Das heißt, es müssten adaptive Systeme entstehen, die beim Grenzübergang ein Update bekommen, beispielsweise zu unterschiedlichen Verhaltensweisen der Verkehrsteilnehmer und gegenüber Verkehrsteilnehmern. Ganz heikel wird es, wenn wir uns darüber unterhalten, dass es Länder geben könnte, wo Frauen weniger wert sind als Männer und Autos dann nur für Männer bremsen. Dann müssten wir uns ernsthaft fragen, ob wir Maschinen wollen,

die das tun. Wir sollten generell Maschinen nicht wollen, die über Leben und Tod von Menschen entscheiden.

Aber sie müssen die Entscheidung treffen ...

Sie müssen die Entscheidung treffen, außer man separiert oder limitiert die Maschinen. Das ist auch die große Debatte. Mein Vorschlag ist es, autonome Autos geschwindigkeitsmäßig zu begrenzen oder sie auf Strecken zu schicken, die gut überschaubar sind. Autobahnen sind dafür optimal.

Menschen bewerten ein autonom fahrendes Auto, das mehr Menschenleben schützt und das Leben des eigenen Fahrers opfert, als moralisch, bevorzugen jedoch für sich selbst das Modell, das seine Insassen am besten schützt – ein Dilemma. Könnte es eine „dosierbare“ Moral geben, bei der der Käufer entscheidet, welche Gewichtung er für sein Auto einstellt?

Sehr schwierig. Autos, die den Fahrer selbst opfern, dürften keine Akzeptanz beim Kauf finden, es sei denn, man wird zum Kauf gezwungen. Die Bundesregierung schafft gerade Fakten. Die Politik will, dass sich der Fahrer mit anderen Sachen beschäftigen kann. Zugleich soll er „wahrnehmungsbereit“ sein und jederzeit eingreifen können – Dobrindt will also autonomes und manuelles Fahren verbinden. Deswegen will er auch das Lenkrad behalten. Das ist klug. Google findet das nicht lustig und wird möglicherweise dagegen klagen.

Es muss eine Harmonisierung und Standardisierung der moralischen Settings geben, zumindest in großen Anwendungsfeldern. Hier kommt noch ein weiteres Problem hinzu: Was Dobrindt jetzt für die deutsche Wirtschaft festlegt, schafft Regeln für die ganze Welt. Das heißt, jetzt beginnt weltweit der große Kampf, was die richtigen Settings, was die richtigen Entscheidungen der Autos sind. Dobrindt übernimmt mit der Ethikkommission aus wirtschaftlichen Gründen eine Vorreiterrolle, denn er spürt den knallharten Angriff aus dem Silicon Valley. Das wird noch riesige Diskussionen auslösen. Meine Grundannahme ist, dass es eine weltweite Harmonisierung in Bezug auf die Settings geben muss. Alles andere wird schwierig.

Reduziert die Algorithmisierung der Welt menschliches Handeln auf maschinell verwertbare Vorgänge? Algorithmen entscheiden, welche Informationen an welcher Stelle innerhalb einer Suchanfrage erscheinen, und sie schätzen ein, ob eine Person kreditwürdig ist. Wieweit kann die Ethik dieser Entwicklung Grenzen setzen?

Hier in der Schweiz wird Predictive Policing angewandt. Es wird eine Software eingesetzt, die Einbrüche und andere Straftaten voraussagt. Ich habe kein Problem damit, wenn es pauschal Straftaten sind, die geschehen sollen, begangen von unbekanntem Übeltätern. Was die Software nicht sagt, ist, dass eine bestimmte Person straffällig wird, wie im Film „Minority Report“. Das wäre sehr heikel. Denn dann sind die Menschen in ihren Persönlichkeitsrechten betroffen. Sobald es individuell wird, habe ich ein Problem damit. Sie nannten die Kreditwürdigkeit. Wohne ich in der falschen Gegend, bin ich nicht mehr kreditwürdig. Deshalb vermeide ich es womöglich, in ein solches Viertel zu ziehen. Es gibt Systeme, die uns beispielsweise an Flughäfen überwachen. Deshalb achten wir darauf, keine Tasche allein zu lassen, weil das System sie sofort erkennen und sie schnell gesprengt würde. Das heißt, wir beginnen uns an die Maschinen anzupassen. Das führt dazu, dass wir uns selbst zensieren. Dabei fing alles mit harmlosen Sachen an, wie sie Amazon eingeführt hat. „Wer das Buch gekauft hat, kaufte auch dieses Buch“ und so weiter; dann fing Netzwerke an, uns Freunde vorzuschlagen. Das ist eigentlich schon sehr merkwürdig, weil ich meine Freunde selber finden will. Inzwischen zieht sich die Schlinge immer weiter zu.

Kann Ethik dieser Entwicklung zur Fremdbestimmung Grenzen setzen?

Wenn wir von der Informations- und der Technikethik reden, kann Ethik hier Reflexionsarbeit leisten. Sie kann die Probleme sichtbar machen. Im Rahmen der Informationsethik schaue ich mir bestimmte Anwendungsgebiete an, wie private Drohnen, Datenbrillen oder Cloud-Computing, und frage danach, welche moralischen Herausforderungen sich beim Einsatz solcher Technologien ergeben. In der Regel tangieren sie unsere Privatsphäre, unsere Persönlichkeitsrechte und unsere informationelle Selbstbestimmung. Das heißt, uns wird etwas aus der Hand genommen, es wird über unseren Kopf entschieden und es werden Daten abgezogen. Es ergeben sich Verschlechterungen aller Art. Genauso sind natürlich Verbesserungen möglich, etwa bei der persönlichen Autonomie.

Dürfen Google und Co. Dinge wissen, die wir nicht einmal unseren besten Freunden anvertrauen würden?

Google aggregiert unter anderem persönliche Daten. Das wurde besonders heikel, als das Unternehmen anfangs, seine verschiedenen Dienste zusammenzuführen. Es weiß von Ihnen Dinge, die Sie vielleicht nur halb bewusst tun. Warum macht Google das? Google ist eine Werbe- und Datenfirma und will Daten sammeln und verwenden. Auch der Einstieg in die Robotik hing stark mit Daten zusammen. Man investiert in Roboter, die Daten absaugen, und damit wären wir wieder bei Servicerobotern oder bei Dingen wie Echo von Amazon oder Hello Barbie. All diese Firmen wollen einerseits Abhängigkeit von Technologien erzeugen und andererseits unsere Daten nutzen. Das nicht unbedingt mit böser Absicht, sondern zu Werbezwecken.

Wie kann ich mich als Einzelner dagegen wehren, dass die Algorithmen mein Leben bestimmen?

Manche Alternativen werden einfach abgeschafft. Sicher kann ich noch Geld am Bankschalter abheben und an einer Kasse bezahlen, aber insgesamt wird es immer schwieriger. Es werden immer mehr Zusammenhänge geschaffen, denen wir uns nicht entziehen können. Das ist das Beängstigende. Man wird aus wirtschaftlichen Interessen Dinge durchsetzen, gegen die wir nicht ankommen werden. Ich versuche relativ konsequent zu sein und habe zum Beispiel keine Kundenkarte, weil ich nicht will, dass das Geschäft weiß, was ich einkaufe. Wenn ich mit Karte zahle, dann werde ich natürlich erkannt. Bargeld bedeutet Freiheit.

Stehen wir dann vor einer Automaten-diktatur?

Ja und nein. Ich glaube nicht, dass Maschinen einmal die Weltherrschaft ergreifen und uns versklaven. Das halte ich für Science-Fiction. Aber wir stehen kurz vor einer schwächeren Form der Maschinendiktatur. Wir machen uns immer abhängiger von Maschinen. Und wir delegieren immer mehr Entscheidungen an sie. Das könnte auch ein Kritikpunkt bei der Disziplin der Maschinenethik sein. Ich bin dagegen, dass man bestimmte weitreichende moralische Entscheidungen einfach Maschinen überlässt. Auf einer grundsätzlicheren Ebene würden wir von Algorithmenethik sprechen. Damit verbunden ist die Frage, was wir an Algorithmen auslagern sollen. Ich würde sagen, nicht alles. Selbst wenn wir Sicherheit gewinnen und Präzision, wir brauchen auch Freiheit, wir müssen Fehler machen können, das gehört zum Menschsein dazu. Wir müssen Geheimnisse für uns behalten können. Deshalb schlage ich mich lieber auf die Seite von Freiheit, Geheimniskrämerei und Fehlerhaftigkeit.

Wie viel Freiheit bleibt uns noch? Können wir sie uns erhalten?

Immer weniger.

Hilft da beispielsweise Wettbewerb?

Ja, das würde helfen. Ich benutze Dienste wie Startpage, die zwar auf der Google-Suche aufbauen, aber meine IP-Adresse nicht weitergeben. Wir haben eine faktische Monopolstellung von Google im Bereich der Suchmaschinen, im Bereich der Maps, auf verschiedenen Gebieten. Es ist ziemlich merkwürdig, dass das so seit zehn Jahren geht. Inzwischen tut die EU ein bisschen was, aber immer noch zu wenig.

Mit steigender Rechenleistung nimmt die KI immer weiter Verbesserungen an unserer Umgebung vor. Ist es eine Frage der Zeit, wann sich die Systeme verselbstständigen?

Autonome Systeme sind hochselbständige Systeme: Sie treffen Entscheidungen unabhängig von uns. Wenn sie miteinander agieren, wie bei der Car-to-Car-Communication, dann entstehen Gebilde, die unabhängig von uns operieren. Was ich nicht glaube, ist, dass Menschen bedroht werden durch Maschinen, die Menschen als Gefahr erkannt haben. Eine kluge Maschine könnte uns vermutlich als die größte Bedrohung für diesen Planeten erkennen. Die Maschinen, die dann selbstständig entscheiden, sie müssten die Menschheit eliminieren, halte ich für unwahrscheinlich, weil wir entsprechende Vorkehrungen getroffen und die erwähnten Metaregeln eingeführt haben.

Lässt sich KI kontrollieren?

Wenn sie relativ primitiv ist, ja; wenn es selbstlernende Systeme sind, wird es immer schwieriger: Wir können Systeme bauen, die sich am Ende des Tages anders verhalten als am Anfang. Ein Beispiel dafür war Tay.

Elon Musk bezeichnet KI als „existenzielle Bedrohung“. Stephen Hawking sagt: „Künstliche Intelligenz könnte das Ende der Menschheit sein.“ Teilen Sie seine Einschätzung?

Eher nicht. Die KI ist eine der spannendsten Forschungstätigkeiten. Wir müssen auch die Gefahren und Risiken benennen können. Die sind aber viel mehr auf einer operativen Ebene vorhanden als auf einer strategischen, globalen Ebene. Viel gefährlicher ist die Maschine, die mich für kreditunwürdig erklärt. Viel gefährlicher sind Systeme, die meine Daten absaugen oder Profile von mir erstellen. Aber dass die KI uns vernichtet, das sehe ich überhaupt nicht. ✕

Goodbot und Liebot

Chatbots auf Websites tun sich schwer damit, richtig auf schwierige Situationen zu reagieren. Oliver Bendel hat deshalb mit Studierenden der Hochschule für Wirtschaft der Fachhochschule Nordwestschweiz den Goodbot entwickelt, einen virtuellen Gesprächspartner, der auf heikle Aussagen und Fragen moralisch angemessen reagieren soll. Der Goodbot hat beispielsweise die Regel, nicht zu lügen, außer wenn er durch das Aussprechen der Wahrheit einen Menschen extrem verletzen würde. Als mögliche Einsatzgebiete sehen die Forscher die Jugendarbeit oder Beratungsangebote von Spitälern oder Kliniken.

Die Wissenschaftler aus der Schweiz haben nun einen Lügenbot entwickelt, der diese Regel komplett umdreht – und gewissermaßen eine unmoralische Maschine geschaffen. Damit soll unter anderem gezeigt werden, dass Maschinen lügen können, und es soll angedeutet werden, wie man ihre Lügen erkennen beziehungsweise wie man ihnen das Lügen abgewöhnen kann. Den Liebot können Sie über luegenbot.ch oder liebot.org testen; die Zugangsdaten können über oliver.bendel@fnw.ch erfragt werden.