



Bauen auf Big Data

Menschen werden vermeintlichem Maschinenwissen ausgesetzt: z. B. Straffällige, deren Rückfälligkeit oder Arbeitslose, deren Vermittelbarkeit mit Künstlicher Intelligenz (KI) berechnet wird.

eine Vermutung – die dritte Stufe. Interessanterweise argumentierte Wundt, dass es sich bei diesem Für-wahr-Gehaltenen noch nicht um ‚Wissen‘ handelt – die vierte Stufe –, sondern zunächst mal nur um eine Meinung bzw. eben Vermutungen.“

Das hieße, dass Datenmärkte mitsamt ihren Big Data-Computern und Künstliche Intelligenzen (KIs) eigentlich nur Meinungen und Vermutungen verkaufen, aber kein Wissen. Darüber sollten sich die Käufer dieser Services im Klaren sein, weil ihnen diese großen Datenanalysen als „Wissen“ verkauft werden. Dass der Wissensbegriff in dieser Weise missbraucht werden kann, hat eine Geschichte. Schon in den 1990er Jahren fingen Firmen an, Datenbanken als „Wissensmanagementsysteme“ zu verkaufen. Viele Firmen fielen bereits damals auf den Trick herein, als sie glaubten, sie könnten das Wissen ihrer Mitarbeiter mit solchen Computersystemen für immer festhalten. Sie stellten dann enttäuscht fest, dass Computer tatsächlich nur Daten speichern und ausgeben können – und all das relevante Lebenswissen der Mitarbeiter in so einem System gar nicht vorhanden sein kann.

Computer haben keine Überzeugungen

Als sich nun die Arbeitsgruppe bei der OECD zusammensetzte (an der ich als Experte beteiligt war), um einen Bericht zu Big Data zu schreiben, wurde dort den großen Datenmengen wiederum die Qualität des „Wissens“ zugeschrieben. Man entschied also an relativ hoher politischer Stelle – und übrigens gegen den Rat einiger beteiligter Industrieexperten, die solche Systeme bauen –, dass Big Data (auf irgendwie magische Weise?) Wissen erzeugt. Doch die Namensgebung von Dingen kann eine ungeheure Macht über die Dinge selbst ausüben. Denn in Folge des Hypes von Big Data kam der Hype von KI, die auf Big Data aufsetzt, und plötzlich wurde auch den KIs Wissen zugeschrieben. Und es wurde vergessen, dass KIs doch gar keine Überzeugung besitzen können, sondern eben nur große Datensätze mit Mustern, Wahrscheinlichkeiten und Korrelationen. Oder haben Sie schon mal einen Computer mit einer Überzeugung gesehen?

Leider hat dieses Vergessen dramatische Folgen für Menschen, die dann in Folge dem vermeintlichen Maschinenwissen ausgesetzt sind: Zum Beispiel Straffällige, deren Rückfälligkeit mit KIs berechnet wird. Oder Arbeitslose, deren Vermittelbarkeit ausgerechnet wird. Oder Schulkinder, deren Schulfähigkeit ausgerechnet wird. Oder Bürger, deren Terrorisierwahrscheinlichkeit ermittelt wird, etc. Alles Beispiele aus dem Leben, hier und heute in Europa. Das vermeintliche „Wissen“ übt also Macht aus und hat damit eine zutiefst ethische Dimension. Eine der wichtigsten ethischen Entscheidungen ist also, Maschinen nicht dort Wissen zuzuschreiben, wo sie nur Vermutungen haben. Wir müssen viel vorsichtiger mit dem Begriff des Wissens umgehen; seine Nutzung hohen Standards unterwerfen. Ganz sicherlich können wir keine Wissensgesellschaft aufbauen, wenn wir leichtfertig mit Sprache umgehen. Wir sind dann keine Wissensgesellschaft, sondern eine verwirrte, unpräzise und ignorante Gesellschaft, die sich mehr einbildet als sie hergibt.

Der vorliegende Beitrag ist die gekürzte Version eines Vortrags am Oscar von Miller-Forum in München.

Die Autorin ist Professorin an der Wirtschaftsuniversität Wien. Ihr Buch „Digitale Ethik. Ein Wertesystem für das 21. Jahrhundert“ ist 2019 im Droemer-Verlag erschienen.

Künstliche Intelligenz setzt heute auf gigantischen Datenmengen auf. Doch es gilt, einen gefährlichen Trugschluss zu vermeiden: Diese Systeme erzeugen kein Wissen, sondern lediglich Vermutungen.

Flüsternder Maschinenklatsch

Von Sarah Spiekermann

Wissen zu entwickeln, ist eines der höchsten Güter unserer Zeit. Dem Laien scheint Wissen auf den allerersten Blick durchaus nur positiv. Aber zu viel wissen zu wollen, vom Baum der Erkenntnis essen zu wollen, war schon immer auch eine Schicksalsfrage für die Menschheit. Zum ersten Mal wurde mir diese eher skeptische Sicht auf Wissen bewusst, als ich dem Vortrag eines Google-Lobbyisten zuhörte, der die Idee vertrat, man könne doch auch eine „Ethik der Ignoranz“ vertreten. Er deutete an, dass man in Unternehmen wie Google, die sehr viel Informationen über Menschen sammeln, oftmals Dinge weiß oder zu wissen meint, die man eigentlich gar nicht wissen will, ja auch gar nicht wissen sollte.

Profile über Onlinenutzer

Das betrifft sicherlich nicht nur das Wissen darum, wie häufig jemand sich Pornoseiten im Internet anschaut, wie oft man beim Eintippen in die Suchmaschine derbe Rechtschreibfehler macht, wie viele Stunden jemand online ist und was er oder sie da alles macht. Das ist ein Wissen der großen Internet-Giganten über uns Nutzer, was nichts mehr ist als eine Vielzahl kleiner trivialer Peinlichkeiten, bei denen man schon fast fragen könnte, ob das Wort „Wissen“ hier das Richtige ist – oder ob man eher von digitalem Klatsch („Gossip“) reden sollte, den sich die Maschinen gegenseitig virtuell zuflüstern, nur um uns dann mit entsprechender Werbung oder Nachhilfe zu überraschen. „Was wissen die schon

von mir“, fragt so mancher Bürger und versucht, die Sache zu verdrängen.

Wirklich große ethische Fragen ergeben sich jedoch, wenn dieser digitale Maschinenklatsch zur Wahrheit erhoben wird. Wir leben in einer Zeit gigantischer Datenmärkte, wo in jedem Moment von uns in Echtzeit circa 30.000 Datenpunkte in so genannten „Personal Data Management Platforms“ zusammengezogen werden. Der Wiener Aktivist Wolfie Christl und ich haben 2017 das Buch „Networks of Control“ herausgebracht mit 600 Fällen, die zeigen, wie Firmen persönliche Daten zusammentragen und vermarkten. Es gibt Firmen, die damit werben, dass sie 700 Millionen Profile über Onlinenutzer weltweit besitzen, mit angeblich 30.000 Attributen pro Person, die sie von 200 Datenzulieferern zusammenziehen.

Solche Daten großer Datenbroker inkludieren in der Regel nicht nur so triviale Dinge wie unser Online-Suchverhalten oder unsere Einkäufe, sondern enthalten – so glauben viele – tatsächlich Wissen über uns. Aus unserem Aufenthaltsort wird er-

„Vermeintliches Wissen über Menschen ist in der Datenmarktwirtschaft zu einer Ressource geworden. Und wir selbst sind zur Erde für diesen Rohstoff geworden.“

rechnet, wo wir als nächstes sein werden oder nicht; aus unseren Stimmprofilen, wie motiviert jemand ist, wie professionell, oder gar welche Persönlichkeit er oder sie angeblich hat. Aus Schrittzählern wiederum wird geschlossen, wie gesund jemand sein muss. Soshana Zuboff hat in ihrem

Buch „Der Überwachungs-Kapitalismus“ eine brillante ökonomische Analyse vorgelegt, die beschreibt, wie sich in den letzten 18 Jahren eine eigene Datenmarktwirtschaft entwickelt hat mit dem Ziel, Wissen über uns zu erwirtschaften, um dieses dann möglichst gewinnbringend weiter zu verkaufen. Vermeintliches Wissen über Menschen ist also zu einer Ressource geworden. Und wir selbst sind zur Erde für diesen Rohstoff geworden.

Aber ist dieser Maschinenklatsch denn wirklich Wissen? Persönliche Daten, Sensordaten und alles, was Computer oder Maschinen über uns erfassen können, sind in Wahrheit nichts mehr als Bits. Es mögen viele Daten sein, aber alle Auswertungen unterliegen letztlich dem Prinzip, dass man sich Häufigkeiten in den Daten ansieht – etwa Korrelationen, Cluster oder Muster. Man versucht dann, mit diesen Daten und ihren Aggregaten Wahrscheinlichkeiten abzuleiten: Mit welcher Wahrscheinlichkeit hängt jemand einer Partei an? Wie wahrscheinlich ist es, dass jemand diese oder jene Krankheit hat? Hat jemand mit hoher Wahrscheinlichkeit betrogen? Ist jemand wahrscheinlich ein Terrorist?

Missbrauchter Wissensbegriff

In meinem Buch „Digitale Ethik“ gibt es folgende Überlegungen zum Wissensbegriff: „In der Antike hat Platon den theoretischen Teil unseres Wissens einmal definiert als ‚eine begründete, wahre Überzeugung‘ von etwas. Alle drei Komponenten dieser Definition – die Begründung, die Wahrheit und die Überzeugung – waren über die letzten 2500 Jahre Gegenstand hitziger philosophischer Debatten (...)

Der Begründer der Psychologie, Wilhelm Wundt (1832–1920), entwickelte eine Herleitung von Wissen, die uns im digitalen Kontext nützlich sein kann. Er unterschied vier Stufen: Glauben, Meinungen, Vermutungen und Wissen. Auf der ersten Stufe reden wir von einem subjektiven Für-wahr-Halten. Das ist der persönliche Glaube an etwas. Auf der zweiten Stufe steht das objektive Für-wahr-Halten, die Meinung. Dieses objektive Für-wahr-Halten stützt sich auf eine ganzheitliche Betrachtung möglichst aller verfügbarer Fakten. Etwas Für-wahr-Gehaltenes sollte also möglichst wissenschaftlich, rational und breit abgesichert sein und gemäß Wundt regelmäßig überprüft werden. Wenn sich solche Meinungen auf etwas Zukünftiges beziehen, handelt es sich um