

ZUM KÜNSTLICHEN UND NATÜRLICHEN WISSEN IN DER MODERNEN ZEIT

Für RPI-Impulse angepasster Auszug aus Sarah Spiekermann:
„Digitale Ethik – Ein Wertesystem für das 21. Jahrhundert“

Nichts ist in der heutigen Zeit wichtiger als die Frage danach, was Wissen ist.

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 zweieinhalbtausend Jahre Gegenstand hitziger philosophischer Debatten, denn was reicht uns als eine Begründung aus, damit etwas als Wissen anerkannt werden kann? Wann ist etwas wahr? Und ist eine Überzeugung dasselbe wie der Glaube an etwas?

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 bzw. Beobachtungen, wie etwa einem weitläufig beobachtbaren und bestätigtem Phänomen, einer wiederholt in dieselbe Richtung weisenden Umfrage oder langfristig bestätigte Ergebnisse von Big Data Analysen. 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 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. Er würde also einem wissenschaftlichen Papier oder einer Big-Data-Prognose den Status des Wissens absprechen. Zu Wissen werden die für wahr gehaltenen Meinungen oder Vermutungen erst dann, wenn Menschen von diesen überzeugt sind. Zwischen einem Für-wahr-Halten und einem Wissen ist also ein qualitativer Sprung, der der Überzeugung bedarf.

Dieser Sprung zwischen Für-wahr-Halten und Wissen ist in seiner Bedeutung nicht zu unterschätzen. Denn

was bedeutet es zum Beispiel bei einer klassischen Debatte, ein Publikum von der eigenen Haltung zu überzeugen? Es geht darum, erstens sinnvolle Argumente zu haben und zweitens unter diesen auch noch die wesentlichen so klug voranzustellen, sodass man am Ende die Wahrheit auf seiner Seite hat. Beide Gegner in so einer Debatte mögen Argumente bringen, die man durchaus für wahr halten kann, aber überzeugen tut nur der, der das meiste Wesentliche und Sinnvolle auf seiner Seite hatte.

Maschinen haben es an dieser Bruchstelle zwischen dem Für-wahr-Halten und dem Wissen nicht leicht, weil sie das Wesentliche und Sinnvolle vom Unwesentlichen oder Unsinnigen schwer unterscheiden können. So könnte ein Kind einen Roboter in der Kirche fragen, warum dort auf dem Altar eine Kerze brennt. Und der Roboter würde durchaus wahrheitsgemäß antworten, dass die Kerze brennt, weil es Sauerstoff in der Luft gibt. Vielleicht könnte er auch noch ausrechnen, dass es wahrscheinlich klüger wäre, zu antworten, dass eine Kerze in einer Kirche Licht spendet oder ein christliches Symbol ist für das Licht in der Welt. Aber warum zünden wir denn eigentlich *wirklich* Kerzen in der Kirche an? Weil Kerzen Wertträger sind. Wenn eine Kerze brennt, dann können sich Werte wie Heiligkeit oder Frömmigkeit,

Prof. Dr. Sarah Spiekermann-Hoff
Leiterin des Instituts Wirtschaftsinformatik und Gesellschaft
an der Wirtschaftsuniversität Wien
mis-sek@wu.ac.at



Erhabenheit, Besinnlichkeit und Lebendigkeit entfalten. Und dass wir uns zu diesen Werten hingezogen fühlen, ist ein ganz maßgeblicher Grund dafür, warum in Kirchen Kerzen aufgestellt werden. Was wir über die Welt wissen, das erschließt sich uns also auch auf einer sinnvollen sinnlichen Wertebene, auf die Maschinen uns nicht folgen können. Wenn man diese vier logischen Stufen der Wissenshierarchie nun akzeptiert, dann unterliegt der Begriff des Wissens einem hohen Anspruch. Man nimmt eine demütige Haltung gegenüber dem Wissensbegriff ein. Wie die Stoiker sagten, bedarf es eigentlich weiser Menschen, die das zu Wissen erheben, was unter dem Für-wahr-Gehaltenen würdig genug ist. Für wahren Fortschritt ist erforderlich, dass nur das zu Wissen erhoben wird, was diese Bezeichnung verdient.

Die Reflexion zeigt, dass es unsere Big-Data-Systeme und KIs schwer haben müssen, einem durchdachten Anspruch auf Wissen gerecht zu werden. Meine bescheidene und stoische Sicht der Dinge ist folglich: Je akkurater, konsistenter, vollständiger, aktueller, größer und kontextsensitiver ein Datensatz ist, desto besser wird die Fähigkeit einer Maschine sein, Muster zu erkennen, die als interessante Maschinenmeinungen oder -vermutungen von uns Menschen herangezogen werden können. Wenn Institutionen nicht weise sind und die Meinungen und Vermutungen von Maschinen pauschal zu Wissen erheben, dann werden selbst sensibelste Urteile in Zukunft auf Maschinen und ihre Datensätze ausgelagert, die in Wahrheit eine Menge Unsinn produzieren können und ganz sicherlich fern jeder Weisheit sind. Aber genau das passiert. Schon in den 1990er Jahren begann die IT-Marketingmaschinerie damit, über ihre Datenbanken und Systeme von »Wissensmanagementsystemen« zu sprechen, und große industriefreundliche Big-Data-Reports, wie der der OECD, erheben zu meinem Leidwesen schon allein *die Analyse* von großen Datenmengen zu Wissen.¹ Schlimmer kann es eigentlich nicht mehr kommen. Man vernebelt den Blick für die erforderlichen Maßnahmen, um sich Wissen anzueignen. Da aber offensichtlich so viel Hoffnung und Herzblut in die Wissensfähigkeit von IT-Systemen gelegt werden, möchte ich der Frage nachgehen und überlegen, welche positive Rolle Maschinen bei unserer Wissensbildung tatsächlich einnehmen könnten, wie sie kraft ihrer speziellen Fähigkeiten zum Fortschritt in diesem Bereich beitragen könnten und wie sie sich hier von uns Menschen unterscheiden.

Menschenwissen versus Maschinenwissen

Der größte Teil unseres Wissens ist nicht theoretischer Natur, sondern umfasst das, was wir im Laufe unseres Lebens internalisieren, beispielsweise wie man einen Garten pflegt. Oder das Wissen, dass jemand ein Kind ist oder ein Greis. Mit der Zeit lernen wir Menschen durch Wiederholung und in der Gemeinschaft mit anderen (etwa unseren Eltern) die Ideen in der Welt kennen – zum Beispiel die Idee vom Alter oder von bewährten Handlungen (beispielsweise dass man für einen Garten immer den richtigen Boden, das richtige Licht und die richtige Wassermenge braucht). Zu diesem Wissensschatz gehören auch die Werte. Wir lernen, wie sich Werte anfühlen: Ordnung, Liebe, Sympathie, Sicherheit. Für unser *implizites* Wissen in Form von Ideen bedarf es keiner Begründung oder eines mathematischen Beweises.

Maschinen kennen diese Art von Wissen im menschlichen Sinne nicht. Sie haben kein Bewusstsein für diese Form von Wissen über das Sein. Sie können Werte nicht wahrnehmen; sie können weder etwas glauben noch von etwas überzeugt sein. Insofern ist eins sicher: Wenn Maschinen intelligent sind, dann sind sie es ganz sicher nicht im menschlichen Sinn. Noch mehr: Weil Computer all die unsichtbaren Ideen und Werte nicht wahrnehmen können, die jeder normal begabte und groß gewordene Mensch zu erkennen gelernt hat, werden uns Computer in dieser Beziehung immer ein bisschen hinterherhinken. All die emotional intelligenten Roboter, wie sie uns in den Filmen *Terminator*, *Ex Machina* oder *Westworld* dargeboten wurden, werden also Science-Fiction bleiben. Es mag Wesen geben, die ähnlich aussehen, aber in ihrem Inneren werden sie anders ticken.

Es besteht die technische Möglichkeit, dass wir Menschen Maschinen Musterbenennungen beibringen, die typischen Ideenausprägungen nahekommen. Zum Beispiel kann man einer Maschine beibringen, dass wir Menschen von einem „Greis“ sprechen, wenn jemand ein Mann ist, weißes Haar hat, viele Falten und gebeugt geht. Oder wir können einer Maschine Gesichtserkennung beibringen. Wir können ihr mitteilen, dass Gesichtsbilder mit hochgezogenen Mundwinkeln und erweiterten Pupillen meistens für ein Lächeln stehen. Solche Vorgänge der Benennung nennt man in den Computerwissenschaften »Tagging« oder auch »Labeling«. Und jeder, der schon mal auf Facebook gewesen ist, kennt diesen Vorgang aus eigener Erfahrung, denn wir werden von Facebook immer wieder zur Angabe aufgefordert, wer auf einem bestimmten Foto zu sehen ist. Eine Künstliche Intelligenz wie die, die Facebook einsetzt, wird so von uns Nutzern in der richtigen Benennung der Dinge und Personen trainiert. Ebenso wie bei Tieren oder Pflanzen übernimmt der Mensch die biblische Aufgabe der Namensgebung. Und so werden Maschinen eine sehr große Menge an »Wissen« aus zweiter Hand von uns über die Welt aufbauen können. Die Ideen selbst aber werden sie nicht erkennen können, denn wenn man plötzlich auf die Idee kommt, eine uralte Katze als „Greis“ zu be-

¹ Ich schreibe hier »zu meinem Leidwesen«, weil ich selbst Teil der Expertengruppe war, die die Entstehung dieses Berichts extern begleitet hat, und wir konnten als Externe den jungen OECD-Mitarbeiter trotz gutem Zureden nicht davon überzeugen, nicht in die Tessa-Falle zu tappen. Er wollte unbedingt Datenanalyse bzw. Verarbeitung selbst schon zu Wissen erheben und hatte die Idee, dass es aus den Daten sozusagen dann auf magische (?) Art zu Wissen käme.

zeichnen, versteht das zwar jeder Mensch, aber eine Maschine nicht.

Maschinen übertreffen andererseits den Menschen beim Erkennen von einmal erfassten und wiederkehrenden Mustern. So kann ein Facebook-Algorithmus beispielsweise Leute auf Fotos wiedererkennen, die wir mit dem bloßen Auge nicht zuordnen können, beispielsweise weil die Haare einer Person auf dem Foto vor dem Gesicht hängen. Künstliche Intelligenz hat ein so feines visuelles Detailerkennungsvermögen, dass sie eine Person sogar auf Basis ihrer einzigartigen Haarstruktur wiedererkennt. Dieser Vorteil kommt auch bei der medizinischen Diagnose zum Tragen, wenn Ultraschallbilder von Krebspatienten analysiert werden oder wenn ein Patient mehrere Medikamente nimmt, deren Bestandteile eine gegenseitige Unverträglichkeit bedingen. In solchen Fällen, wenn hoch strukturierbare, wiederkehrende Muster in großen Datenmengen auftreten, sind Maschinen einfach großartig. Sie erkennen etwas, was wir Menschen zwar theoretisch wissen, aber selbst nicht so leicht sehen können. Sie erspüren Wissenswertes für uns.

Maschinen können auch Realitäten in einer Vollständigkeit sehen, die uns Menschen entgeht. In der Forschung spricht man vom Phänomen der Unaufmerksamkeitsblindheit beim Menschen, um zu beschreiben, dass wir oft nur das sehen, worauf wir unsere Aufmerksamkeit fokussieren oder was wir erwarten und sehen wollen. Was nicht unseren Denk- und Wissenschaftsmodellen oder unseren Erwartungen entspricht, wollen wir nicht wahrhaben. Hier sind (vor allem selbstlernende) Maschinen hilfreich. Weil eine Künstliche Intelligenz eine gegebene Realität wie ein zunächst gleichgewichtetes Pixelspace voller Datenpunkte begreift, kann sie Muster und Feinheiten erkennen, die uns Menschen entgehen. Wenn wir sie auf großen Datenmengen über Sozialsysteme trainieren, dann kann es auch sein, dass sie Muster erkennt, auf die wir selbst noch nicht gekommen wären, die aber einen Wahrheitsgehalt haben. Sie stellt sozusagen Vermutungen an, die sich weiterverfolgen lassen, um vielleicht neues Wissen zu generieren.

Ein Kollege, den ich in der Big-Data-Arbeitsgruppe der OECD getroffen hatte, der selbst ein Marketing-Intelligence-Unternehmen in Paris gegründet hat, erzählte mir, dass die Big-Data-Analysen in seiner Firma aufgedeckt hätten, dass der Vorname einer Person mit der wichtigste Vorhersagefaktor im Kaufverhalten sei. Zwar gibt es schon in der jüdischen Kabbala die Idee, dass unsere Namen sehr viel über uns aussagen, dass eine Sarah nun mal keine Susi ist, aber im Prinzip hat die Maschine hier einen Zusammenhang in den Daten entdeckt, die viele moderne Denker als esoterisch abtun würden. Wenn Maschinen solche Schlussfolgerungen anstellen und diese auf Basis von qualitativ hochwertigen Daten mit geringer Fehlerwahrscheinlichkeit begründen können, dann erscheinen sie uns intelligent, weil sie unsere eigenen Hypothesen und unser Wissen über die Welt infrage stellen und bereichern. Sie machen auch mit dem klassischen aber problematischen Modellproblem der Moderne Schluss, weil sie komplemen-

tär zu unseren bestehenden Theorien manchmal neue Muster erkennen, die relevant sein können, oder in großen Datenmengen Zusammenhänge ermitteln, die wir nicht erwarten würden und daher kaum sehen können (manchmal auch nicht wahrhaben wollen). Wenn diese relevant sind, können sie zu neuem Wissen führen. Das ist das wahre Fortschrittspotenzial Künstlicher Intelligenzen. Und ich sehe sehr viel digitales Fortschrittspotenzial für unsere Wissenschaften, wenn die Kooperation von Menschen und Maschinen so geartet ist, dass wissende Menschen unerwartete Maschinenvorschläge und Mustererkennungen unter die Lupe nehmen.

Was wir darüber aber nie vergessen sollten ist, dass Maschinen gleichzeitig keinen natürlichen Zugang haben zum Reich der Werte und damit zum Reich der Bedeutung. Sie haben eben eine „künstliche“ Intelligenz, die von der „natürlichen“ Intelligenz des Menschen sehr verschieden ist. Und wer das nicht beachtet und nicht versteht, der wird Künstlichen Intelligenzen eine Weisheit zusprechen, die sie systemisch niemals besitzen können. Es wird ein Vertrauen geschenkt werden, was nur enttäuscht werden kann.

