DIE WELT ALS FLUGHAFEN? DATENBANKEN ALS MYTHEN UND BEDROHUNG

Ein amerikanischer Supermarkt sammelte die Einkaufsdaten eines Mädchens. Noch bevor der Vater des Mädchens über dessen Schwangerschaft informiert war, wussten die Daten schon davon und die Coupons für Babyartikel waren bereits unterwegs. Ein Aufschrei der Empörung ging durch die Medien: Der Kunde – ein transparentes Wesen! Kontrolliert von Algorithmen! Der Autor Felix Keller entlarvt diese Geschichte als Angstmacherei und relativiert die Macht gesammelter Daten.

Keywords: Datenbank, Kaufverhalten, Beobachtung, Algorithmen, Nicht-Orte, Big Data, Neuromarketing

Felix Keller

Jedes Zeitalter erzählt seine Legenden. Jüngst raschelte eine Geschichte durch den (digitalen) Blätterwald, die so plausibel erschien, dass sie sich ohne weiteres Nachforschen, ohne weitere Überlegungen nacherzählen liess: In der Nähe von Minneapolis stürmte ein aufgebrachter Vater in eine Filiale der Supermarktkette Target und beklagte sich beim Manager, seine Tochter hätte Coupons für Babywaren erhalten. Seine Tochter sei minderjährig und keinesfalls schwanger. Ob Target kleine Mädchen zur Mutterschaft animieren wolle? Der Manager entschuldigte sich und meldete sich einige Tage später nochmals telefonisch, um sein Bedauern zu bekräftigen. Zerknirscht antwortete der Vater, seine Tochter hätte sich wohl Aktivitäten hingegeben, von denen er nichts gewusst habe. Sie sei, so habe er nun erfahren müssen, tatsächlich in Erwartung. Des Rätsels Lösung: Der Supermarkt sei, so berichtet der Blätterwald, im Besitz eines Algorithmus, der das Kaufverhalten der Tochter ausgewertet und so ihre Schwangerschaft korrekt diagnostiziert habe. Das Datenbankprogramm wusste mehr über die Tochter als deren Vater. Der Name und das Logo der Supermarktkette, eine stilisierte Zielscheibe, erscheinen nun beinahe emblematisch: Das Ziel der neuen Macht der Daten und der Mathematik sind die Einzelperson und ihr Privates. Gnadenlos jagen die Algorithmen den Einzelnen aufgrund der Spuren, die er notgedrungen hinterlässt, hoffnungslos ist der private Mensch einer neuen Sichtbarkeit ausgesetzt. Die Geschichte um die schwangere Tochter und den unwissenden Vater gelte dahingehend bereits als «klassischer Fall», weiss ein Nachrichtenmagazin.[1] Ein «klassischer Fall» – aber wovon genau?

Bezahlen, einpacken und ab in die Datenbank

Es lohnt sich, das Zustandekommen dieser sich viral verbreitenden Geschichte genauer zu betrachten. Die Quelle dieser Erzählung lässt sich, obwohl kaum genannt, unschwer rückverfolgen: Es handelt sich um einen Artikel in der *New York Times*, datiert auf den Februar 2012 mit dem Titel *How Companies Learn Your Secrets*.[2] Der Titel, mit der direkten Ansprache des Lesers, zielt auf die Wahrnehmung einer von Konzernen durchleuchteten Privatsphäre. Verfasst ist der Text von Charles Duhigg, der am Schluss des Artikels das baldige Erscheinen seines eigenen Buchs zum Thema ankündigt. Im Wesentlichen beschreibt Duhigg gängige neuropsychologische Untersuchungen zum Zustandekommen alltäglicher Gewohnheiten, die zumeist automatisch und reflexionslos ausgeführt werden. Die allgegenwärtigen Routinen haben ihren Sinn: Sie entlasten die kognitiven Kapazitäten und sparen Aufmerksamkeitsressourcen ein. Es existiert nun ein aufstrebender Wissenschaftszweig des Neuromarketings mit dem Ziel, in diese Gewohnheiten gezielt einzugreifen. Die Modifikation von Habits sei dann besonders gut möglich, wenn eine so drastische Umstellung des Lebens ansteht wie eine Mutterschaft.

Hier tritt der Held der erwähnten Legende in Aktion: Andrew Pole. Er hat Wirtschaft und Statistik studiert und preist sich selbst als «math nerd». Anhand eines Registers, in das sich die werdenden Mütter bei Target eintragen können, um Vergünstigungen für ihre «Baby Shower»-Party zu erhalten,[3] konnte Pole mit einer statistischen Analyse ungefähr 25 Produkte identifizieren, die werdende Mütter gehäuft einkaufen. Aufgrund dieser Basisinformationen konstruierte er einen Punkte-Index, der ihm erlaubte, jeder Kundin von Target einen «pregnancy prediction score» zuzuweisen. Soweit ist dies ein bekanntes methodisches Vorgehen: Die Wahrscheinlichkeit einer Schwangerschaft kovariiert mit der Zahl der Indexpunkten, die sich aus den gekauften Produkten ergeben. Auf der Basis der nationalen Datenbank von Target spuckte

1 http://www.profil.at/articles/1310/560/354256/zahlenmaessig-big-data.

2 http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html.

3 http://www.target.com/c/baby-registry/-/N-564tn

der entsprechende Algorithmus eine Liste von Zehntausenden von Frauen aus, «who were most likely pregnant». Wie präzise sein Modell ist erläuterte er nun mit eben der Geschichte der vorhergesagten Schwangerschaft, die dann um die Welt ging und als Beleg für die Macht von Algorithmen und Big Data herhalten musste.

Entwarnung – Algorithmen sind auch nur Zahlen

Der Autor des Artikels fragt nichts nach – ob es beispielsweise weitere Fälle gibt oder auf welche Weise die Reliabilität der Schätzung geprüft werde. Aber gerade in diesem Beispiel, in dem alle zeitgenössischen Erwartungen und Ängste angesichts systematischer Datenauswertung geballt vorkommen, steckt die Krux des Wissens der Statistik: nämlich in einer Leerstelle zwischen Daten und konkretem Fall. Denn die Aussage, dass eine konkrete Person mit dieser oder jener Wahrscheinlichkeit schwanger ist, ist aufgrund der Datenanalyse schlicht nicht möglich. Die Analyse lässt nur Aussagen über Kategorien der Population zu, auf denen die wahrscheinlichkeitstheoretische Berechnungen basieren, aber nichts über den konkreten Einzelfall (genau so wenig wie sich die Wahrscheinlichkeit der Augenzahl des konkreten nächsten Würfelwurfs bestimmen lässt). Mehr noch, für das so stolz präsentierte «Ergebnis» ist eigentlich die Analyse überflüssig. Nehmen wir an, den Zehntausenden von Frauen, die in Targets Datenbank figurieren und im gebärfähigen Alter sind, werde hälftig nach dem Zufallsprinzip Coupons für Babyware zugestellt beziehungsweise nicht. Der Hälfte der Gruppe von Frauen in der Datenbank, die schwanger sind und deren näheren Umgebung noch nichts von der Schwangerschaft weiss, fände also Babycoupons im Briefkasten – und dies dürften angesichts der Zehntausenden von Frauen in der fraglichen Situation nicht wenige sein. Irgendwo würde mit aller Wahrscheinlichkeit ein misstrauischer Vater einer pubertären Tochter, die ganz und gar zufälligerweise in das Sample mit Babycoupons gefallen war, in Rage geraten.

Mit anderen Worten gesagt: der von den Medien kolportierte Fall hätte sich genau so gut ereignen können, hätte Target eine zufallsgenerierte Streusendung verwendet. Aber die Empörung, die hunderttausendfache Diskussion im Internet aber auch in seriösen Medien wie *Die Zeit* und *Der Spiegel, New York Times* und *Forbes* zeigt, dass die Erzählung, mit der ein junger Mensch seine eigene Fähigkeiten anpreist, auf einen Nährboden gefallen ist, auf dem sie zu einer urban tale wurde. Sie hat sich mit Hilfe der neuen Medien viral fortgepflanzt und so die Botschaft verbreitet: Dein privates Leben ist heillos exponiert, geheime Algorithmen wissen alles über dich! Die Erzählung, so gering ihr konkreter Gehalt auch sein mag, trifft auf den Nährboden einer Zeit, in der eine Bedrohung des Privaten durch die immer grössere Berechenbarkeit durch die Auswertung immer grösserer Datenmengen erkannt wird, die jemand unweigerlich in seinem Leben hinterlässt. Freilich bedeuten mehr Daten nicht notwendig mehr Wissen: Der Berechenbarkeit und der Vorhersage sind enge theoretische Grenzen gesetzt, die auch ein noch so raffinierter Algorithmus nicht zu überwinden vermag.[4]

Die Rückkehr des aussengeleiteten Menschen

Doch die Bedeutung der Erzählung um Targets Beobachtungs- und Prognosefähigkeit und ihrer hunderttausendfachen medialen Multiplikation liegt weniger in der angeblichen Genauigkeit der Berechnung, sondern vielmehr darin, dass sie die Wahrnehmung einer umfassenden Bedrohungslage ausdrückt, die so transportiert und vervielfacht wird und damit, ungeachtet des Wahrheitsgehaltes, wahre Konsequenzen erzeugt. «Wenn Menschen Situationen als real definieren, so haben sie reale Konsequenzen», formulierte der Soziologe William I. Thomas dieses Theorem. Hinsichtlich dieser durchaus realen Konsequenzen einer imaginären Bedrohung ist jüngst die sehr aufschlussreiche Metapher der Welt als Flughafen aufgetaucht: der Flughafen als Ort der Überwachung und Kontrolle schlechthin. Niemand würde einen Witz über eine Bombe in einem Flughafen machen – jeder versucht sich so unauffällig wie möglich zu verhalten, um am besten durch die Kontrolle zu kommen, um ja nicht die Aufmerksamkeit des Sicherheitspersonals auf sich zu lenken, so der Blogger, der die Metapher zuerst ins Spiel brachte.[5]

Die umfassenden Möglichkeiten des Beobachtetseins regeln das Verhalten, ohne dass effektiv der Akt einer Beobachtung geschieht. Ich vermeide auffälliges, nicht-normales Verhalten weil ich registriert werden könnte – überall dort, wo ich beobachtet werden kann, also aufgrund der umfassenden Nutzung elektronischer Geräte und Daten auch im privaten Leben. Auf eine neue Weise scheint der «aussengeleitete Mensch» aufzuerstehen, von dem David Riesman in den 1950er-Jahren in seinem Buch «Die einsame Masse» berichtete: Jener Mensch, der sich perfekt an die soziale Normen hält und dadurch jede Individualität verliert betrachtete Riesman als Typus des modernen Menschen schlechthin. Heutzutage sind es nicht mehr die generalisierten «Anderen», von denen man sich beobachtet und kontrolliert fühlt, sondern ein vermeintlich sehr reales, ein technologisches Dispositiv, das sich stillschweigend immer weiter ausbreitet und immer mehr Bereiche durchdringt. Der Effekt bleibt gleich wie in der Enge der Gesellschaft der 1950er-Jahre: nicht auffallen, normal sein.

4 Darauf kann hier leider nicht weiter eingegangen werden. Die ersten Enttäuschungen über die Versprechen der Prognosefähigkeit von Big Data-Analysen setzen ein. Vgl. hierzu: Butlar (2013), sowie Lazer (2014). Nate Silver (2013) zeigt, weshalb der Anteil der Fehlprognosen in der Big Data-Ära sogar zunehmen wird.

5 https://medium.com/i-m-h-o/9a1e5268ff39

Der Nicht-Ort als Ort der Passage…

Aber dass Riesman das Argument schon in den 1950er-Jahren verwendet hatte, zeigt, dass die technische Beobachtung Bestandteil einer viel grösseren Transformation ist. Der Flughafen wie der Supermarkt lassen sich als paradigmatische Orte moderner Gesellschaften begreifen. Der französische Ethnologe Marc Augé hat diese im Gegensatz zu anthropologischen Orten, an denen alles seinen Sinn und seine Geschichte hat, als Nicht-Orte beschrieben (beispielsweise Dorfplatz vs. Bahnhofsunterführung einer Grossstadt). Sie zeichnen sich dadurch aus, dass in ihnen einander unbekannte Menschen ohne einen Bezug zueinander zirkulieren. Es sind Orte der Passagen und damit der Anonymität. Mit anderen Worten gesagt: Diese Orte haben keinen anderen Sinn, als dass sie passiert werden und man nur einen Moment darin verweilt. Sie erzeugen kein soziales Milieu, in dem sich Dinge, Menschen und Zeichen zu einer einigermassen dauerhaften Ordnung fügen. Sie sind symbolisch unterdeterminiert. Es herrscht zwar eine überdeutliche Ordnung der Zeichen: Schilde, Signaturen und Hinweise, die einem ans Ziel führen. Ein Supermarkt ist sogar ein eigentliches Reich von Signaturen, doch die Menschen gleiten gleichsam unter diesen Zeichen hindurch, da sie in dieser Ordnung nur temporär anwesend sind.

… und als Ort der heillosen Überforderung für die Datensammler

Je höher die Geschwindigkeit und je grösser der Umfang, mit denen Menschen, Güter und Daten zirkulieren, desto eher transformieren sich anthropologische Orte, in denen Dinge und Person symbolisch einen zugewiesenen Platz werden, zu solchen Nicht-Orten. Und je grösser der Blickwinkel jener Instanzen wird, die die Ordnung beobachten, desto eher erscheinen Milieus als Nicht-Orte, in der die zirkulierenden Menschen als schwierig fassbar erscheinen; als ephemere Gestalten. Die Gesellschaftswissenschaften haben diesen Prozess beobachtet und mit Hilfe von Konzepten wie *disembedding* (Entbettung), Entstrukturierung oder Auflösung geschlossener Milieus definiert. Wie sie auch genannt wird, diese Transformation hinterlässt aus der Beobachterperspektive ein Defizit in einer symbolischen Ordnung, die immer auch Kontrolle war: eine selbstverständliche Erkennbarkeit von Individuen in Raum und Zeit – jedes hatte seinen konkreten Platz. Der Versuch, das «Individuum» wieder als konkret eindeutigen Punkt, als Adresse in diese Orte einzufügen, bedeutet einen immensen Aufwand an Technologien der Beobachtung und Adressierung, der Erfassung und Kontrolle. Vielleicht wäre es fruchtbar, Big Data, das grenzenlose Datensammeln sowie die Programmierung scheinbar mächtiger Algorithmen nicht nur als Erfolgsgeschichte, als «Megatrend», sondern auch als ein Symptom der heillosen Überforderung zu betrachten, als mehr oder weniger sinnvolle Aktion, doch noch irgend etwas zu tun, angesichts einer Realität, die zu entschwinden droht, Terabyte um Terabyte an Daten anhäufen, in der vergeblichen Hoffnung, dass irgend einmal ein Algorithmus Wissen hervorbringt.

Der Philosoph Gilles Deleuze spricht in einem 1990 veröffentlichten, beinahe prophetisch anmutenden Text davon, wie sich das Individuum, das Unteilbare, angesichts der Auflösung geschlossener Milieus (die nicht nur die romantisch verklärte Bürgerstube umfasste, sondern auch Institutionen wie Schule oder Disziplinar-Anstalten) zum «Dividuum» wandelt; zu einem Geteilten, das gleichzeitig an mehreren Orten existiert und ebenso beobachtet werden kann – nicht zuletzt aufgrund der neuen Technologien: An wie vielen Orten ist einer genau, wenn er an der Bushaltestelle soziale Netzwerke bedient und mit der Kreditkarte das Billett löst? Durch diesen Prozess sind Scheidelinien wie «privat» und «öffentlich», zwischen denen Individuen hin und wechseln wie zwischen Kaffeehaus und bürgerlicher Stadtwohnung, schon längst zerfallen. Urbane Legenden wie die hier erwähnte nisten genau in diese Leerstelle, die den Verlust dieser einstmals klaren und bedeutenden Kategorien gesellschaftlicher Organisation bezeichnet, ohne dass sich neue Konzepte abzeichnen.

Wenn die Daten unlesbar werden…

Gibt es keinen Ausweg aus diesem Teufelskreis von Beobachtung und Selbstbewahrheitung? Vielleicht gibt die Art und Weise, wie professionelle Identitätsveränderer und -verstecker heute vorgehen, Anhaltspunkte: Sie löschen eben nicht Daten, sie verstecken eben nicht Informationen über eine Person an geheimen Orten, die sowieso bald durch ein neues Sicherheitsleck der ganzen Welt offen stehen, sondern sie fügen im Gegenteil immer mehr und immer neue Daten und Spuren hinzu, verquere und richtige, so dass nicht mehr klar ist, welche authentisch sind und welche nicht. Misinformation, Disinformation, Digital Distortion nennt sich das (vgl. Ahearn 2010). Die Spurenverwischer handeln getreu dem Diktum, das Roland Barthes schon vor langer Zeit formuliert hat, nämlich dass die beste Subversion eines Codes nicht darin liegt, ihn zu zerstören, sondern ihn zu entstellen. Das Unlesbarmachen, Verändern, Überschreiben, Multiplizieren bis zur Unübersichtlichkeit verhindert die Möglichkeit, im Dividuellen das Individuelle und Adressierbare zu fixieren. Dies bedeutete, die Privatsphäre ganz neu zu denken – oder die Idee am Besten ganz aufzugeben. Ganz trivial und alltäglich: Die Pflege der eigenen Unberechenbarkeit wäre Bestandteil einer neuen Sorge um sich selbst wie um andere.



Felix Keller

Felix Keller lehrte und forschte nach dem Studium der Soziologie und politischen Philosophie an den Universitäten Zürich, Lausanne, Luzern und Fribourg. Zurzeit arbeitet er als Assistenzprofessor an der Universität St. Gallen. Er beschäftigt sich vornehmlich mit wissenssoziologischen Fragen der Vermessung und Sichtbarmachung gesellschaftlicher Realitäten sowie der Dynamik gesellschaftlicher Utopien.

Literatur:

Augé, Marc (1994): Orte und Nicht-Orte. Vorüberlegungen zu einer Ethnologie der Einsamkeit. Frankfurt am Main: Fischer.

Ahearn, Frank M./Horan, Eileen C. (2010): How to disappear: Erase your digital footprint, leave false trails, and vanish without a trace. Guilford, CT: Lyons Press.

Butlar, Declan (2013): When Google got flu wrong. US outbreak foxes a leading web-based method for tracking seasonal flu, in: Nature, 494, S. 155-156.

Lazer, David (2014): The Parable of Google Flu: Traps in Big Data Analysis, in: Science, 343, S. 1203-1205.
Silver, Nate (2013): Die Berechnung der Zukunft. München: Heyne.