

Das Erkennen von Mustern und die Grenzen der Messbarkeit
SWR2 Radio Akademie: Der vermessene Mensch (12)

Autor: Maximilian Schönherr

Redaktion: Detlef Clas

Regie: Günter Maurer

Sendung: Samstag, 19. Juli 2014, 8.30 Uhr, SWR2 Wissen

Bitte beachten Sie:

Das Manuskript ist ausschließlich zum persönlichen, privaten Gebrauch bestimmt. Jede weitere Vervielfältigung und Verbreitung bedarf der ausdrücklichen Genehmigung des Urhebers bzw. des SWR.

Mitschnitte auf CD von allen Sendungen der Redaktion SWR2 Wissen/Aula (Montag bis Sonntag 8.30 bis 9.00 Uhr) sind beim SWR Mitschnittdienst in Baden-Baden für 12,50 € erhältlich.

Bestellmöglichkeiten: 07221/929-26030

SWR 2 Wissen können Sie ab sofort auch als Live-Stream hören im SWR 2 Webradio unter www.swr2.de oder als Podcast nachhören:

<http://www1.swr.de/podcast/xml/swr2/wissen.xml>

Manuskripte für E-Book-Reader

E-Books, digitale Bücher, sind derzeit voll im Trend. Ab sofort gibt es auch die Manuskripte von SWR2 Wissen als E-Books für mobile Endgeräte im so genannten EPUB-Format. Sie benötigen ein geeignetes Endgerät und eine entsprechende "App" oder Software zum Lesen der Dokumente. Für das iPhone oder das iPad gibt es z.B. die kostenlose App "iBooks", für die Android-Plattform den in der Basisversion kostenlosen Moon-Reader. Für Webbrowser wie z.B. Firefox gibt es auch so genannte Addons oder Plugins zum Betrachten von E-Books.

<http://www1.swr.de/epub/swr2/wissen.xml>

Kennen Sie schon das Serviceangebot des Kulturradios SWR2?

Mit der kostenlosen SWR2 Kulturkarte können Sie zu ermäßigten Eintrittspreisen Veranstaltungen des SWR2 und seiner vielen Kulturpartner im Sendegebiet besuchen.

Mit dem Infoheft SWR2 Kulturservice sind Sie stets über SWR2 und die zahlreichen Veranstaltungen im SWR2-Kulturpartner-Netz informiert.

Jetzt anmelden unter 07221/300 200 oder swr2.de

Dieses Manuskript enthält Textpassagen in [Klammern], die aus Zeitgründen in der ausgestrahlten Sendung gekürzt wurden.

MANUSKRIFT

Radio Akademie Intro: Der vermessene Mensch

Ansage:

Das Erkennen von Mustern und die Grenzen der Messbarkeit
Von Maximilian Schönherr

Erzähler:

Neugierig war der Mensch schon immer. Er zieht gern aus Beobachtungen Schlüsse. Er wirft hundert Mal einen Stein, zieht daraus den Schluss, dass die Steine immer eine Parabelbahn durchfliegen – und baut dreitausend Jahre später ein Ding, das verdammt gut trifft, auch über Kontinente hinweg, und nennt es Boden-Boden-Rakete.

Wir Menschen sortieren Beobachtungen zu Mustern. Der Nachbar guckt immer böse, also will er uns Böses. In anderen Ländern wird dann gleich das Gewehr bereitgelegt, bei uns die Mäuler der üblen Nachrede gespitzt und die Fäuste geballt. Auch wenn der Nachbar nur unter Depressionen leidet und ganz harmlos ist und unseren guten Zuspruch gebrauchen könnte.

Messen und Muster erkennen, wie es heute Sensoren und Computer in großen Mengen tun, war in der Geschichte der Wissenschaft schon Mitte des 19. Jahrhunderts üblich, damals natürlich mit Lineal, Papier und Feder. Hier beschreibt der Dresdner Kriminologe Erich Wulffen Vermessungsarbeiten seines Kollegen Theodor von Bischoff. Sie dienten stets dazu, Gemeinsamkeiten zu erkennen, Muster; in diesem Fall: aus der Kopfgeometrie zu ermitteln, ob ein Mann zum Verbrechen neigt oder nicht:

Sprecher 1:

Untersuchte Verbrecherschädel: 206. 23,1 Prozent sind asymmetrisch, bei Geisteskranken beträgt die Asymmetrie nur 7 Prozent, bei Normalen 3,7 Prozent. Starkes Hervortreten der Augenbrauenbogen und der Stirnhöhlen bei 58,3 Prozent der untersuchten Verbrecherschädel. Erinnert an Wilde und prähistorische Menschen, Neandertal-Schädel, mit deutlichem Atavismus. Die mittlere Hinterhauptsgrube vorhanden in 16 Prozent der männlichen Verbrecher, 3,4 Prozent der weiblichen Verbrecher, 4,1 Prozent bei normalen Männern, in 3,2 Prozent bei normalen Weibern, in 26,0 Prozent bei Wilden und in 14 Prozent bei Irren. Unter 11 Fällen mit Hinterhauptsgrube sind 6 Diebe und 5 Mörder.

Erzähler:

Diese Verfahren scheiterten in den wissenschaftlichen Kreisen von Medizin und Psychiatrie einige Jahrzehnte später. Um das Jahr 1900 herum war man sich weitgehend einig, dass sich aus körperlichen Eigenschaften kein Verbrechertyp herauskristallisieren lässt. Aber die Grundidee blieb latent im kollektiven Bewusstsein. Die Nazis nutzten sie als praktisches Mittel, Juden und psychisch Kranke zu typisieren und zu vernichten.

Heute hat jeder Banker so mächtige statistische Werkzeuge an der Hand, dass er unter Millionen von Wertpapieren dasjenige herausfiltern kann, welches mit einer bestimmten Wahrscheinlichkeit seinen Wert innerhalb der nächsten halben Stunde verzehnfachen wird. Die Software erkennt Muster. 2002 war das noch Science-Fiction.

Sprecher 1:

Polizist 1:

Sprecher 2:

Kommen Sie auf Händen und Knien aus dem Schrank!

Erzähler:

Frühes Drehbuchmanuskript für Minority Report von John Cohen, 1997.

Sprecher 1:

Nichts passiert. Dann sprengen die Polizisten mit kurzen Ultraschall-Schüssen die Schranktür.

Erzähler:

In dem Film „Minority Report“ von Steven Spielberg, mit Tom Cruise in der Hauptrolle, kommt ein Staat an sehr viele sensible Informationen über seine Bürger heran und missbraucht diese zu Unterdrückungszwecken.

Sprecher 1:

Polizist 1:

Sprecher 2:

Auf Händen und Knien!

Sprecher 1:

Johnny kriecht auf die Polizisten zu, hebt seinen Kopf. Johnny:

Sprecher 2:

Ich hab nichts getan!

Erzähler:

Das Argument ist ganz modern: Schutz vor Verbrechen. Die Polizei kann nämlich mithilfe dieser Daten Gewaltverbrechen vorhersehen und den zukünftigen Tätern schon vorher das Handwerk legen.

Sprecher 1:

Polizist 2 sieht auf einem Gerät die Worte „Identifikations-Scan“ und zielt mit einem roten Laser in Johnnys Augen, um die Iris zu identifizieren. Der Polizist nickt dem anderen bestätigend zu. Polizist 2:

Sprecher 2:

John Palmer, positiv.

Sprecher 1:

Polizist 1 (zu Johnny):

Sprecher 2:

John Palmer, ich verhafte Sie als den zukünftigen Mörder Ihrer Frau Carol Palmer. Sie haben keine Rechte.

Erzähler:

Der Preis für die auf 100 Prozent gestiegene innere Sicherheit ist die totale Kontrolle der Bevölkerung.

„Minority Report“ steht in einer langen Tradition von Literatur, in der das Sammeln von Informationen über Bürger zum brutalen Machtinstrument wird – wie in George Orwells „1984“, wo ein nie sichtbarer „Big Brother“ alles sieht. Erschienen 1949.

Es gibt seit dem Aufkommen von Personal Computern und Internet zahlreiche Fälle von Datenverstößen. Aber erst die vom Whistleblower Edward Snowden 2013 aufgedeckte NSA-Überwachung gießt die großen negativen Utopien der Literatur in unsere heutige Wirklichkeit: Ein Geheimdienst überwacht pauschal Telefongespräche und Internetverbindungen weltweit. Niemand ist davor sicher, nicht einmal Staatschefs wie die deutsche Bundeskanzlerin.

Offizieller Zweck: Terrorvorhaben gegen die USA und ihre Verbündeten schon in der Planung zu ersticken.

Cut 1: Keith Alexander

Sprecher 1 (= Übersetzer):

Die heutigen Kommunikationssysteme sind die ausgefeiltesten, die die Menschheit je entwickelt hat. 2,5 Milliarden Personen kommunizieren weltweit über dieses geniale Netz.

Terroristen, Kriminelle, Feinde bedienen sich derselben Infrastruktur. Sie verstecken sich darin. Sie nutzen das, was alle nutzen: Gmail, Facebook, Twitter usw. Die moderne Technik hat es ihnen leicht gemacht.

Erzähler:

Die Rede von General Keith Alexander, Leiter des Geheimdiensts NSA, vor dem US-Kongress Ende 2013, enthält, so harmlos sie klingt, all den Stoff, aus dem Orwells Roman und Spielbergs Film gemacht sind.

Cut 2: Keith Alexander

Sprecher 1 (= Übersetzer):

Wir müssen diese Technik in den Griff bekommen, um Schädlinge auszumachen und unschuldige Bürger zu schützen.

Erzähler:

„Unschuldige Bürger schützen“, sagt der General.

Cut 3: Keith Alexander

Sprecher 1 (= Übersetzer):

Wir sind uns bewusst, dass unsere Analysesoftware ein mächtiges Instrument in den Händen des Staats ist. Aber wir nutzen es nur im gesetzlich erlaubten Rahmen. Wir schützen damit unsere Nation.

Erzähler:

Und, so Keith Alexander, wenn einmal was schiefgeht, ist es ein Versehen und von geringer Bedeutung. Das braucht man nicht so hochzukochen wie Snowden es tat.

Cut 4: Keith Alexander

Übersetzer. Sprecher 1:

Uns unterlaufen dabei auch Fehler. Die sind aber unbeabsichtigt und kommen in allen komplexen technischen Systemen vor. Die in der Presse bekannt gegebenen Tausenden von Fällen, wo die Privatsphäre verletzt wurde, sind falsch und irreführend.

Erzähler:

Die National Security Agency NSA rechnet der massenhaften Überwachung keine konkreten Zahlen und Fälle gegen, bei denen sie Anschläge oder Verbrechen verhindert hätte. Sie sagt, als Geheimdienst muss sie das nicht.

Im Wochentakt veröffentlichten Zeitungen wie der englische *Guardian* neue von Edward Snowden gesammelte Details – so viele, dass sie trotz ihrer politischen Bedeutung inzwischen nur noch dritte Plätze in den anderen Medien einnehmen. Nur ein Beispiel von vielen: Ende Februar 2014 veröffentlichte der *Guardian* ein von Snowden bezogenes Papier, demzufolge der britische Geheimdienst GCHQ mithilfe der amerikanischen National Security Agency Webcam-Bilder aus den Wohnzimmern von Millionen von Internetnutzern abgefangen und gespeichert hat.

[Das Programm zum Mitschneiden von Einzelbildern privater PC-Kameras heißt „Optic Nerve“ – Optischer Nerv. Allein in einem Zeitraum von 6 Monaten im Jahr 2008 hat der Geheimdienst durch das heimliche Überwachen von Internetleitungen Fotos von weltweit mehr als 1,8 Millionen Yahoo-Nutzern mitgeschnitten.]

Im streng geheimen, GCHQ-internen Originaldokument heißt es:

Sprecher 1:

Wegen der großen Datenmenge greift die Software nur alle 5 Minuten ein Einzelbild ab. Es wäre hilfreich, ein Werkzeug zu haben, das viele Bilder schon vor dem Abspeichern auswählt; dann könnten wir Chatfotos in kürzerer Folge sammeln.

[Erzähler:

Zudem beklagt der britische Geheimdienst in dem Dokument, dass unter den heimlich gesammelten Fotos 7 Prozent Nacktfotos seien – von ermittlungstechnisch geringem Wert. Die Software, die Bilder mit viel nackter Haut aussondern soll, funktioniert nicht zuverlässig. Zudem lenkten die vielen anzüglichen Webcam-Fotos die Geheimdienstmitarbeiter von der eigentlichen Terroraufklärungsarbeit ab.

Ein Analysewerkzeug von „Optic Nerve“ nutzt ein Verfahren, das wir auch von normalen PCs und Macintosh-Computern kennen: die biometrische Analyse von Gesichtern, also zum Beispiel den Abstand der Augen oder die Form der Ohren. Am PC zu Hause dient das dazu, die Oma, die man auf einem Foto „Oma“ genannt, also verschlagwortet hat, automatisch auch in hundert anderen Fotos als Oma zu erkennen. Man findet die Fotos dann sofort, wenn man nach „Oma“ sucht. Klappt nicht immer, aber häufig.

Dem Geheimdienst ist die Fehleranfälligkeit dieser digitalen Kopfvermessung bekannt, weswegen ein Abschnitt in dem Papier erwägt, die höher auflösenden Webcam-Bilder der Microsoft X-Box-Computerspielekonsole abzufangen; Gamer spielen heute meistens mit anderen übers Internet. Und sie haben dabei ihre Kamera eingeschaltet, etwa um ihren Krieger mit Gesten, also mit Körperbewegungen zu steuern.

Und die geheimdienstlichen Begehrlichkeiten gehen noch weiter:]

Sprecher 1:

„Wir sollten uns auch mit der Iris-Erkennung beschäftigen. Denken Sie an Tom Cruise in *Minority Report*.“

Erzähler:

Wörtliches Zitat aus dem Papier des britischen Geheimdienstes.

Dagegen wirken Datenschutzthemen von früher fast romantisch, als manche sich darüber aufregten, wenn der Online-Buchhändler Amazon ihm aufgrund des Kaufs eines Buches ähnliche Bücher anbot. Oder der Computerkonzern Apple, der gern alle unsere Bilder und Musik in der Cloud, also in einem Rechenzentrum irgendwo, abspeichern möchte. Keine Sorge, so Apple, ...

Sprecher 1:

In der Cloud sind Ihre Daten sicher. Und es ist für Sie sehr praktisch, denn Sie können die Daten von jedem Gerät weltweit jederzeit abrufen!

Erzähler:

Der Suchmaschinenbetreiber Google geht noch weiter und fordert den Kunden dazu auf, gespeicherte Fotos scannen zu dürfen. Damit sind Algorithmen gemeint, die die Bilder nach Tag/Nacht, nach Urlaub/Arbeit, nach alt/jung und dergleichen analysieren. Zu angeblich anonymisierten, rein statistischen Zwecken. Harmlos?

Der Datenhunger von Google ist ungebremst. „Street View“-Autos fahren durch Städte und Dörfer und fotografieren jedes Haus. „Google Glass“, eine Brille mit Internet und eingebauter Kamera, nimmt unbemerkt alles auf, was ihr Träger sieht. Der Widerstand dagegen, insbesondere in der Öffentlichkeit, wo sich Menschen unangenehm beobachtet fühlen, nimmt zu.

[Sprecher 1:

Wenn Ihnen das passiert, nehmen Sie die Datenbrille bitte ab, drehen Sie sie nach hinten oder stecken Sie sie in Ihre Tasche ...

Erzähler:

... rät Google in einer neuen Richtlinie zur Benutzung der Datenbrille ...

Sprecher 1:

Und lesen Sie auf dem ins Brillenglas eingebauten Bildschirm nicht Schmöker wie Krieg und Frieden. Man könnte Sie für verrückt halten.]

Erzähler:

Der NSA-Skandal war auch für die Wissenschaft ein Paradigmenwechsel. Zuvor dachten die meisten Informatiker: Eine solche Datenmenge, wie sie beim Abfangen von E-Mails, Telefonaten und Videochats entsteht, kann niemand auswerten. Es ist einfach zu viel, es sprengt jeden Speicher. Sie suggerierten, dass wir uns sicher fühlen können in dieser gigantischen Datenmenge, einfach weil sie so gigantisch ist. Sinngemäß: Lass die Geheimdienste ruhig sammeln, sie werden selbst daran ersticken.

Seit Edward Snowden die internen Mechanismen offengelegt hat, sehen auch Informatiker das etwas anders. Die NSA baut gerade ein neues Rechenzentrum, so groß wie ein Dorf, mit großem Datenspeicher, viel Software – und Supercomputern. Es scheint doch möglich zu sein, mit statistischen Verfahren, Mustererkennung, Entschlüsselungstechnik große Datenmengen auf sogenannte Signifikanzen – Besonderheiten, Auffälligkeiten – zu untersuchen. Ob die Ergebnisse einen Sinn haben, können wir nur vermuten, sonst würde die „Dienste“ nicht so massiv in Hard- und Software investieren.

Akustischer Trenner

Erzähler:

Das Sammeln und Verknüpfen von Daten ist heute so negativ belegt, dass wir leicht aus dem Blick verlieren, wie nützlich es sein kann, z.B. in der Spitzenforschung, wo große Datenmengen und Mustererkennungen ihren Sinn haben.

Cut 5: Dimitri Kanevsky

There can be hundreds of scenarios where it can be useful ...

Sprecher 1 (= Übersetzer):

Wir können uns Hunderte an Szenarios für unsere Technik vorstellen.

Erzähler:

Der russisch-amerikanische Forscher Dimitri Kanevsky. Er ist gehörlos. Ihm sind über 100 Patente zu verdanken, fast alle auf dem Gebiet der digitalen Stimmanalyse.

Cut 6: Dimitri Kanevsky**Sprecher 1 (= Übersetzer):**

Stellen Sie sich ein Baby vor, das nachts häufig weint. Den Eltern wurde gesagt, nicht jedes Mal sofort zu springen, wenn das Kind weint, weil es dadurch verwöhnt wird. Besser fünf oder auch zehn Minuten warten.

Gut: Die Eltern hören ihr Töchterchen im Kinderzimmer nebenan weinen und machen sich Sorgen. Sollen wir rübergehen oder nicht? Vielleicht braucht die Kleine sofortige Zuwendung?

Unser System versteht, warum das Baby weint, und hilft den Eltern dabei, eine kompetente Entscheidung zu treffen. Wir können die Laute des Babys analysieren und den Eltern signalisieren: Da ist nichts passiert, ihr braucht jetzt nicht sofort zu ihr rüberzugehen. Oder in einer anderen Situation: Euer Baby hat Bauchschmerzen und braucht eure wärmende Hand. Unser System erkennt, wenn das Kind sich vor etwas fürchtet und Sie es beruhigen und trösten müssen.

Erzähler:

Der Datenschutz spielt bei Dimitri Kanevskys Forschung keine Rolle, aber die Grenzen der Analyse liegen auf der Hand. Noch kann er nicht mit so hoher Sicherheit sagen, dass die Software das Schreien des Kinds richtig interpretiert, um es in die Praxis umzusetzen. Aber es funktioniert so gut, dass es über die Grundlagenforschung hinaus ist.

Cut 7: Dimitri Kanevsky

When we teach a computer ...

Übersetzer. Sprecher 1

Wir müssen uns daran orientieren, wie unser Gehirn lernt. Unser Hirn hat ja gelernt, zahllose Muster wahrzunehmen und zu interpretieren. Es hat eine erstaunliche Rechenkapazität, die auf einer sehr großen Menge an Neuronen beruht, die alle miteinander verbunden sind.

Unser kognitives Computersystem imitiert deswegen die kognitive Rechnerarchitektur des Gehirns. Damit kann es sehr differenzierte Muster im Schreien von Kindern erlernen.

Erzähler:

Bei der Datenanalyse gibt es zwei Richtungen. Die eine rechnet nach einem starren, vorher festgelegten Schema – sogenannten deterministischen Algorithmen –, Datenberge durch und sucht nach Verbindungen. Damit das zu Ergebnissen führt, muss man sehr präzise Fragen stellen, etwa: „Bei wie vielen Babyschreien sind Frequenzen von über 10.000 Hertz im Spiel?“ Dimitri Kanevsky und andere Wissenschaftler in einem Zukunftsprojekt bei IBM gehen einen völlig anderen Weg. Ihre Computersysteme sind lernfähig, man kann mit unscharfen Fragen an sie herantreten. Sie funktionieren wie neuronale Netze, so wie das Gehirn ein Netz von Neuronen, Nervenzellen ist.

Cut 8: Hendrik Hamann

Gucken Sie sich einmal unser Gehirn an: Wie funktioniert es? Man weiß es nicht genau, aber eine Sache wissen wir: Wir haben unendlich viele Prozessoren, die ganz langsam sind. Und wir haben Speicherkapazität direkt neben dem Sensor. Das heißt, es ist ein massiv paralleles System, das ganz langsam läuft.

Sie brauchen diese neuen Computer, weil die Sensoren so viele komplexe Informationen liefern?

Richtig, weil die sehr komplexe Informationen liefern. Aber auch, weil die Sensoren viel mehr Informationen *verbinden* müssen. Und dafür muss viel mehr paralleles Prozessieren von Informationen stattfinden. Und der dritte Grund ist natürlich, dass der Computer *lernen* soll. Und um wirklich gute neuronale Netze zu bauen, wenn Sie solche parallelen Computer haben, ist es natürlich viel einfacher, die sind sehr viel stärker.

Erzähler:

Hendrik Hamann arbeitet wie Dimitri Kanevsky im IBM-Projekt „5 in 5“ – fünf (menschliche) Sinne in fünf Jahren.

Es gibt viele Projekte weltweit, die die Welt als Sensor-Landschaft begreifen, die Signale in Form von Babygeschrei, von Bild, von Temperatur liefert. Paul Lukowicz nutzt in einem europäischen Forschungsvorhaben namens „FuturICT“ das Smartphone als Sensor.

Cut 9: Paul Lukowicz

Es gibt in London etwas, das nennt sich „Lord Mayor's Show“. Das ist ein großer Karneval, da sind eine halbe Million Leute auf den Straßen, und es ziehen verschiedene Wagen von Vereinen durch, also wie Karneval.

Wir haben jetzt eine App für die Mobiltelefone der Leute geschrieben. Diese App gibt ihnen Informationen über die Show, wie sie dahin kommen können, wo welche Verkaufsstände sind. Gleichzeitig haben wir die Leute gebeten – das müssen sie bestätigen, und sie können das auswählen –, dass sie uns gerne zu Forschungszwecken periodisch ihre Location, also ihre Koordinaten geben, die wir dann anonymisieren.

Und daraus haben wir dann so eine Art grafische Darstellung von der Dichte der Menschenmenge und wie sich die Menschenmenge bildet erzeugt. Der Sinn besteht darin, dass man zum Beispiel pro-aktiv solche Dinge, wie sie, glaube ich, in Duisburg [gemeint ist die Katastrophe bei der Love Parade im Juli 2010 mit 21 Toten] passiert sind, erkennen und verhindern kann. Aber es ist eben die Frage, wie kann man diese Daten von den Handys effektiv sammeln, und wie kann ich daraus Aussagen oder Vorhersagen darüber ableiten, was sich da in der Menschenmenge tut?

Cut 10: Philipp Heim

Es ist in gewisser Weise eine Erweiterung der Sensorik, die wir bekommen.

Erzähler:

... sagt der Informatiker Philipp Heim.

Cut 11: Philipp Heim

Wir haben einen riesigen Raum um uns, die Welt, unsere Gesellschaft, die Städte, die Infrastruktur, die bevölkert ist mit Menschen. Und diese Menschen offerieren uns jetzt quasi über das Web 2.0, wenn sie denn mobile Endgeräte haben und gewillt sind, Informationen über ihre Umwelt preiszugeben, laufend das, was um sie herum passiert. Zum Beispiel der Amoklauf in Winnenden: In diesem konkreten Fall war es so, dass sich der Täter innerhalb eines Gebäudes befand. Das heißt, die Polizei konnte nicht einsehen, was in dem Gebäude vor sich ging, sie hatte keine „Sensoren“, die ihnen die Informationen lieferten, die sie gebraucht hätten.

Ich kann auch nicht einfach in das Gebäude hineingehen, weil ich Leute dadurch gefährde, dadurch, dass ich nicht weiß, was ich damit auslöse. In welcher Situation befinden sich jetzt gerade die Schüler?

Auf der anderen Seite war es wirklich der Fall, dass die Schüler, wie sie sich in den Klassenräumen verschanzt hatten, über Web 2.0 Informationen zur Tatzeit abgaben, über die Lage vor Ort. Das heißt ganz konkret: „Auf dem Flur vor unserem Klassenraum wird geschossen.“ Oder: „Der Täter bewegt sich weg.“ Oder: „Wir sind hier im Klassenraum XY und haben uns eingeschlossen.“

Solche Informationen sind im konkreten Fall unheimlich wichtig, um auf Seiten der Polizei z.B. die richtigen Entscheidungen zu treffen. Gerade Bilder und Videos, Youtube und Flickr, bieten hier einen enormen Mehrwert, gerade auch für Feuerwehrmänner, die nicht auf die Beschreibung möglicherweise eines Laien angewiesen sind, der jetzt „einen Rauch“ sieht und ihn beschreibt mit „hier brennt's, hier brennt's!“ Jetzt richtig einzuklassifizieren, um welche Brandstufe es sich hier handelt.

Wenn ich jetzt aber ein Bild von der aktuellen Situation vor Ort habe, dann kann ich als Experte eine Einschätzung treffen. Und das ist ein Riesenvorteil gegenüber einem textuell abgesetzten Notruf.

Wie kann ich es als Einsatzleiter schaffen, in diesen riesigen Web 2.0-Datenmengen genau die Informationen herauszubekommen, um die Situation vor Ort so einzuschätzen, dass ich die richtigen Aktionen treffe und Schaden vermeide?

Erzähler:

Das was für die akademische Forschung naheliegt und fast paradiesisch klingt – Daten für einen guten Zweck herzugeben –, kann natürlich jederzeit auch gegen uns genutzt werden. Unzählige Anwendungen auf Smartphones rufen, ohne uns zu fragen, die GPS-Daten ab, also unseren Standort. Oder sie aktivieren bei Bedarf, ohne uns zu fragen, das Mikrofon oder die Kamera. Wohin diese Daten dann gelangen, liegt im Nebel. Sie landen zunächst natürlich im Rechenzentrum des App-Herstellers, der selbstverständlich in seinen Allgemeinen Geschäftsbedingungen schreibt, dass er sie anonymisiert und niemals weitergibt. Dem kann man nicht vertrauen, denn viele Hersteller solcher kleinen kostenlosen Programme haben Verträge mit Werbetreibenden, und sie selbst wissen in der Regel nicht einmal, wo das Rechenzentrum steht, in dem die Daten ihrer Kunden liegen. Was passiert, wenn das Start-Up-Unternehmen einmal verkauft wird oder pleitegeht? Angenommen, Google würde Insolvenz anmelden, in 50 Jahren, aus Gründen, die wir nicht kennen: Die Rechenzentren werden dann sicher nicht gesprengt, im Gegenteil.

Doch wir müssen nicht *nur* pessimistisch sein. Der Forscher für Künstliche Intelligenz, Stefan Wrobel:

Cut 12: Stefan Wrobel

Die Daten in sozialen Netzwerken sind hochinteressant, weil sie einen Bereich unserer Gegenwart, unserer Wirklichkeit beleuchten, der bis jetzt eher wenig beleuchtet war. Wir haben schon lange in der Wissenschaft Messgrößen aufgenommen und festgehalten. Wir haben schon lange im Bereich des Geschäfts Zahlenprozesse abgebildet: Einnahmen, Ausgaben, Kaufprozesse. Aber wir hatten lange Zeit kein in irgendeiner Weise präzises Abbild der sozialen Struktur unserer Gesellschaft, dessen, was Menschen zusammenhält, was sie miteinander verbindet.

Wenn Sie da zurückblicken auf soziologische oder auch psychologische Studien – vor 10 oder 20 Jahren konnte man froh sein, wenn man diese Studien über einige Dutzende oder einige Hundert Leute durchführen konnte.

Heute bilden soziale Netzwerke auch diesen Teil unserer Realität zumindest ansatzweise ab, sodass das natürlich hochinteressante Informationen sind, aus jedem Blickwinkel.

Für die empirische Forschung sind das unschätzbare Werte, die da entstehen, weil wir viel mehr über den Zusammenhalt der Gesellschaft lernen können. Sie bieten aber auch enorme Potenziale für Dienstleistungen, für Produkte, für Computer, die uns besser unterstützen können, bei dem, was wir im Alltag tun wollen. Denn je mehr der Computer grundsätzlich darüber weiß, was ich tun möchte, desto besser kann natürlich auch die Unterstützung sein.

Erzähler:

Die Daten der sozialen Netzwerke wie Facebook liegen vor uns, Statistikprogramme gibt es kostenlos und kostenpflichtig in großer Zahl. Ein solches Programm dient speziell der Auswertung von Versteigerungen über das Internetportal Ebay.

Viele Ebayer wurden deswegen Kunden der Statistikfirma, weil sie alles in ihrem Keller und auf ihrem Speicher versteigert hatten, es ihnen irgendwie Spaß machte, aber dann war Stillstand: Was verkaufe ich jetzt, was wollen die Leute haben? Genau darauf gibt das Software-Werkzeug Antwort. Es listet zum Beispiel Produkte auf, die von den Besuchern der Webseite besonders häufig nachgefragt wurden, und dann kann man die Software fragen: Welche dieser Anfragen ging ins Leere, weil das Produkt nicht oder nur selten angeboten wurde?

Das Programm liefert auch detaillierte Grafiken des zeitlichen Verlaufs zig Tausender von Versteigerungen. Daran sieht man zum Beispiel, dass viele Auktionen mitten in der Woche zu Ende gehen und dann deutlich niedrigere Preise erzielen als solche, die zum Beispiel am Samstagnachmittag enden. Man sieht auch, dass Wodka beim britischen Ebay-Portal viel teurer ist als bei Ebay Frankreich. Dass mit einem jetzt nicht mehr topaktuellen Smartphone der Spitzenklasse allein im deutschen Ebay-Store täglich um die 60.000 € umgesetzt werden.

Sprecher 1:

„Finden Sie Ihre Konkurrenten, analysieren Sie deren Erfolg!“

Erzähler:

... wirbt das Programm. Wie wär's mit günstig in Taiwan Festplatten einkaufen und hier versteigern? Da lässt man besser die Finger davon, denn wie will man gegen den Händler, der in den letzten drei Monaten etwa eine Million Euro mit rund 10.000 Festplatten umgesetzt hat, antreten?

[Was die Software nicht sagt, muss man zu Fuß erledigen, sich den Verkäufer also direkt ansehen, und da findet man neben guten Bewertungen in den letzten Wochen zahlreiche schlechte und Kommentare dieser Art:

Sprecher 1:

Ich möchte den Verkäufer melden, der scheinbar in betrügerischer Absicht tätig ist. Nachdem ich den Kaufpreis für eine 3 Terabyte-Festplatte überwiesen hatte, hatte ich dem Kauf dann doch widersprochen innerhalb der Widerspruchsfrist – mein Bauchgefühl bei diesem Händler war denn doch einfach zu schlecht. Noch immer ein Geldeingang. Da von diesem Anbieter plötzlich auch keine Angebote mehr zu finden sind, gehe ich von einem Betrüger aus.]

Erzähler:

Computerprogramme haben ihre Grenzen. Auch wenn sich die mathematischen Disziplinen der Wahrscheinlichkeitstheorie und Statistik immer weiterentwickelt haben, braucht jedes Szenario sein eigenes Modell. Christian Bluhm, Banker und Finanzmathematiker, über den Alltag eines Finanzmaklers, englisch: „Quant“.

Cut 13: Christian Bluhm

Was die Quants machen: Der fängt vielleicht um 8 oder so an, manche schlafen ein bisschen länger. Dann fahren sie ihre Rechner hoch, öffnen ihre Software; das kann Mathematica sein, Matlab, SPSS, S+, R sein, da gibt es ganz viel Software, und dann fangen sie an, an dem Problem, an dem sie am Abend zuvor aufgehört haben, weiterzubasteln.

Typisches Problem?

Typisches Problem wäre, Sie haben 10.000 Firmenkundenbilanzen und müssen aus diesen 10.000 Firmenkundenbilanzen ein System erstellen, das Ihnen sagt, welcher Kunde risikoreicher als ein anderer ist. Das tun Sie, indem Sie die Bilanzkennzahlen auswerten, man sagt dazu: ein Scoring-System bauen, ein Ranking. Sie haben am Schluss dann die Vorstellung, dass der 5897. Kunde ein bisschen risikoärmer als der eine Nummer davor, aber ein bisschen mehr als der eins danach ist.

Das System blickt in die Vergangenheit ...

... blickt in die Vergangenheit, ergänzt es aber oft noch mit zukünftigen Prognosen wie makroökonomische Faktoren. Diese Analyse, von der ich gerade gesprochen habe, passiert in der Regel nicht nur in *einer* Softwareumgebung, sondern Sie haben eine statistische Software, oft eine sehr analytische Software wie Matlab und Mathematica sowie Datenbankenprogramme.

Ich würde sagen, die Hälfte der Zeit verbringt ein Quant eigentlich mit dem Aufbereiten der Daten. Sie müssen über SQL-Datenbanken auf das Datawarehouse der Bank zugreifen. Das ist nicht alles Mathematik. Ich würde sagen, die Formellastigkeit eines Quant-Jobs liegt vielleicht bei 30 bis 40 Prozent, der Rest ist mühsame Datenarbeit und das Auswerten, Aufbereiten usw.

Erzähler:

Ein Beispiel, bei dem die mathematischen Modelle versagt haben, ist die Bankenkrise um das Jahr 2009. Die Finanzmathematikerin Claudia Klüppelberg kritisiert, dass die Banken sich nicht der neueren Mathematik bedient haben, sondern den bequemeren Weg gingen und mit gewohnten Algorithmen Finanzentwicklungen simulierten:

Cut 14: Claudia Klüppelberg

Die mathematische Theorie, die hinter diesen Abhängigkeiten steht, nennt sich Extremwerttheorie. Das ist sicherlich auch ein Problem, dass in den Banken die Quants häufig von dieser Theorie noch überhaupt nichts gehört haben.

Cut 15: Christian Bluhm

Es gibt sicher in konservativeren Instituten eine gewisse Ablehnung gegenüber neomodischen Modellen. Es kommt auch ein bisschen darauf an, ob sich Ihre ganz oberen Chefs darauf einlassen.

Erzähler:

Wir leben in einer spannenden Zeit, wo uns die Daten zu Füßen liegen und wir mit den richtigen Werkzeugen die Welt besser machen können. Aber überall lauert die Gefahr, dass Dritte in der Lage sind, sich unerlaubten Zugang zu unseren zu Daten verschaffen. Die Konsequenzen von beidem kann man nicht abschätzen. Hauptsache, so Paul Lukowicz vom Deutschen Forschungszentrum für Künstliche Intelligenz, wir sind uns über unseren Schatten bewusst:

Cut 16: Paul Lukowicz

Ich spreche gerne von dem sogenannten „digitalen Schatten“. Heutzutage ist es so, dass jedes Ereignis in der reellen Welt einen Schatten, ein Echo in der digitalen Welt hinterlässt. Und die Frage, wie man diesen Schatten interpretiert, ist schwierig. Ich wähle den Vergleich mit dem Schatten mit Absicht, denn wenn Sie sich überlegen: Jedes Objekt hat einen Schatten, aber der ist mal klein, mal groß. Ein Schatten kann eine sehr gute Wiedergabe der Wirklichkeit sein, er kann aber auch wahnsinnig täuschen. Sie können in einem Schattentheater zwei Hände für einen Drachen halten.

* * * * *