

# Die gesteuerte Zukunft

Die Macht der Algorithmen erreicht die nächste Stufe: Unternehmen, Behörden und Wissenschaftler versuchen mit Hilfe gigantischer Mengen persönlicher Daten, unser Leben zu lenken. Wollen wir das?



ZUMA PRESS / ACTION PRESS

An lauen Frühlingsabenden bietet der Blick von der Hamburger Köhlbrandbrücke auf den Hafen eine Postkartenidylle: Die Elbe glitzert im Abendrot, Stapler, Kräne und Lastwagen bewegen sich wie in Zeitlupe, ab und an gleitet ein Containerschiff vorbei. Aus der Sicht von Sebastian Saxe ist das Areal aber auch dann vor allem eine Gleichung mit vielen Unbekannten. Der 57-Jährige ist Mathematiker und arbeitet seit viereinhalb Jahren bei der Hafengesellschaft an seiner bislang kniffligsten Rechenaufgabe.

Das Hafengebiet ist 7200 Hektar groß. Etwa 200 Züge rollen täglich über die 300 Kilometer Schienennetz, um die angelandeten Güter über 130 Brücken abzutransportieren. Saxe soll diesen logistischen Alptraum optimieren. Es geht um viel bei dieser Mission des Chief Information Officer der Hamburger Port Authority (HPA).

Die Fläche ist begrenzt, ein Ausbau kaum noch möglich. Dennoch hat der Hamburger Senat das Ziel ausgerufen, den Containerumschlag in Hamburg bis 2025 fast zu verdreifachen. Das wird nur gehen, wenn es Saxe und seinem 60-köpfigen IT-Team gelingt, eine andere Ressource optimal auszuschöpfen: Daten. Und davon hat er jede Menge.

Heute schon wimmelt es im Hafen von Sensoren: Lkw und Güterzüge funken ständig ihre Position, einlaufende Containerschiffe ihren Standort und ihre Geschwindigkeit. Auch in der Köhlbrandbrücke sind Sensoren verbaut, die ständig den Verkehr überwachen.

„Unser Ziel ist der total vernetzte, intelligente Hafen, der Smart Port“, sagt Saxe. Ein Hafen, in dem eine Eisenbahnklappbrücke sich beispielsweise nicht mehr zu einer bestimmten Uhrzeit öffnet, sondern kurz bevor das Schiff sie tatsächlich erreicht – wodurch die unnötigen folgenden Verspätungen bei der Bahn und am Terminal vermieden werden. Sogar die Köhlbrandbrücke soll „intelligent“ werden, indem sie über die Sensoren ihren aktuellen Zustand meldet und den nächsten Wartungstermin vorhersagt. Der Sanierungstakt musste gerade verkürzt werden, weil deutlich mehr Schwertransporte darüber verkehren als bislang angenommen. Das interessiert nicht nur Saxe und die HPA, sondern auch Zoll und Polizei, denn ein Teil der Transporte war illegal.

So soll aus der komplexen Hafeningenieuristik letztlich eine Maschine werden, die sich selbst steuert. Saxes Zukunftsvorstellung ist eine Art Hafenbörse, auf der Spediteure auf die Minute genau vorausberechnen können, wie schnell ihr Container vom Wasser auf die Straße kommt.

So geht es gerade vielen Unternehmen weltweit: Sie entdecken einen Rohstoff neu, den sie selbst, ihre Anlagen und Kunden im Überfluss produzieren: Daten.

# 56



## PROZENT

### MEHR FESTNAHMEN

**GELANGEN DER POLIZEI VON  
SANTA CRUZ, KALIFORNIEN, DURCH  
DEN EINSATZ VON KOMPLEXEN  
DATENKOMBINATIONEN.**



US-Polizisten in Santa Cruz

DAN COVRO / PICTURE ALLIANCE / DPA

Die Fachwelt hat sich auf den Begriff „Big Data“ geeinigt, um den Trend zu beschreiben, der in den Vorstandsetagen, auf Kongressen wie der Berliner re:publica vorige Woche und in etlichen neuen Büchern so heiß gehandelt wird.

„Big Brother“ war gestern. Nun heißt Big Data zwar auch totale Kontrolle, verspricht dafür aber zudem eine sinnvolle Steuerung unserer Zukunft in allen Lebenslagen. Autoren wie der Oxford-Professor Victor Mayer-Schönberger sprechen von einer „Revolution“. Big Data, so betitelt er sein aktuelles Buch zum Thema, werde unsere Arbeitswelt und sogar unser Denken verändern.

Entscheidend ist dabei nicht die schiere Datenmenge, auch wenn sie derzeit mit einer Wucht anschwillt, die alles Bisherige in den Schatten stellt: 2012 sollen welt-

# 50



**PROZENT**

**DER POTENTIELL VERWERTBAREN  
DATEN STAMMTEN IM JAHR 2010 AUS**

**ÜBERWACHUNGS-  
KAMERAS.**

QUELLE: IDC

weit 2,8 Zettabyte an Datenvolumen entstanden sein – das entspricht einer Zahl mit 22 Stellen. Experten prognostizieren, dass es bis 2020 sogar 40 Zettabyte werden könnten. Allein im Internet werden tagtäglich so viele Daten herumgeschickt, dass rund 250 Millionen DVDs gebraucht würden, wollte man sie alle abspeichern. Und auch dieses Volumen verdoppelt sich annähernd alle zwei Jahre.

Wirklich neu ist, wie Unternehmen, Behörden und Wissenschaftler ihre Datensätze nun auszuwerten und zu analysieren beginnen. Weil Speicherplatz fast nichts mehr kostet und die Rechner immer schneller werden, können sie inzwischen rund um die Uhr unterschiedlichste Datenarten verknüpfen und korrelieren. Algorithmen sind die Spürhunde in diesem virtuellen Raum. Sie durchwühlen das Chaos, entdecken ungeahnte Muster – und enthüllen so zugleich neue Zusammenhänge, Erkenntnisse und Geschäftsmodelle.

Und auch wenn die wenigsten Bundesbürger bislang etwas mit Big Data anfangen können, erleben sie die neue Macht der Algorithmen schon allerorten: Kreditkartenfirmen verwenden die Analysen, um ungewöhnliche Nutzungsmuster schnell zu erkennen – und Karteninhaber automatisch zu warnen, wenn mit ihrer Kreditkarte plötzlich irgendwo große Rechnungen beglichen werden, obwohl sie bislang nie vor Ort waren. Energiefirmen analysieren den idealen Standort für Windräder mit Wetterdatenanalysen bis auf den Meter genau. In der schwedischen

Hauptstadt Stockholm werden Algorithmen zum Verkehrsmanagement eingesetzt – laut offiziellen Angaben haben sich die Fahrzeiten durch die Innenstadt seither halbiert, die Emissionen sind um zehn Prozent zurückgegangen. Und Online-Händler nutzen die Analysen längst, um ihre Verkaufsstrategien zu optimieren. Das weitverbreitete „Kunden, die diesen Artikel kauften, interessierten sich auch für ...“ ist dafür nur ein Beispiel.

Google und Facebook sind Big Data in Reinkultur. Ihre Geschäftsmodelle basieren auf dem Sammeln, Auswerten und Vermarkten der Informationen ihrer Nutzer – durch möglichst passgenau zugeschnittene Werbung. Den Facebook-Anlegern war diese gigantische Datenbank und die Ahnung, was man im Big-Data-Zeitalter mit mehr als einer Milliarde individueller Profile anstellen kann, gut hundert Milliarden Dollar wert.

Mittlerweile beflügelt die Aussicht, den eigenen Datenschatz zu Dollars zu machen, die Phantasie vieler Branchen, von Supermärkten über die Automobilindustrie und Luftfahrt bis zu Banken und Versicherungen. 2012 lag der weltweite Umsatz für Big-Data-Anwendungen nach Zahlen des Branchenverbands Bitkom bei 4,6 Milliarden Euro. Bis 2016 soll er auf knapp 16 Milliarden steigen.

Auch in der Medizin und der Wissenschaft generell wird mit zahllosen Big-Data-Ansätzen experimentiert. Selbst die öffentliche Hand und besonders Polizei und Sicherheitsbehörden, sonst in IT-Fragen nicht eben vorneweg, haben die Potentiale für die eigene Arbeit erkannt.

Was viele elektrisiert, ist die Verheißung, dank der blitzschnellen Massen-Datenanalysen in die Zukunft schauen zu können: Tatsächlich lassen Algorithmen offenbar erstaunlich präzise Vorhersagen über menschliches Verhalten zu – sei es vor dem Supermarktregal oder im Straßenverkehr bis hin zur Zahlungsmoral bei fälligen Kreditraten.

Schon 2010 sagte Google nur aufgrund des Suchverhaltens seiner Nutzer eine Grippewelle voraus. Der amerikanische Datenspezialist Nate Silver prognostizierte das Ergebnis der jüngsten US-Präsidentschaftswahlen lange im Voraus und präziser als alle Demoskopien.

Sogar die Wahrscheinlichkeit von Verbrechen in bestimmten Stadtteilen wird in einigen Städten vorausberechnet. „Predictive Policing“ heißt das Verfahren, vorausschauende Polizeiarbeit, die an eine Hollywood-Fiktion erinnert: In Steven Spielbergs „Minority Report“ wurden einst Täter für Verbrechen verhaftet, die sie noch gar nicht begangen hatten.

Auch die mutmaßlichen Delinquenten ausfindig zu machen ist offenbar kein Problem: Wissenschaftler haben herausgefunden, dass sich mit Hilfe unserer Handy-Ortungs- und Adressbuchdaten

ziemlich sicher vorhersagen lässt, wo wir uns morgen oder in einem Jahr zu einem bestimmten Zeitpunkt aufhalten werden.

Weil dank Big Data unser Leben und Verhalten berechenbar zu werden scheint, spricht der amerikanische Tech-Guru Chris Anderson vom „Ende der Theorie“. Zu ähnlichen Befunden kommt der österreichische Medienmanager Rudi Klausnitzer, der sein aktuelles Buch zum Thema „Das Ende des Zufalls“ genannt hat.

Warum sollten wir uns künftig auf die Intuition verlassen, wenn alles dank Datenanalyse vorhersagbar zu werden scheint? Und wo bleiben in so einer Welt die Überraschungen, die Fortschritt erst provozieren?

Tatsächlich verlassen sich viele schon jetzt ausgerechnet im intimsten Lebensbereich auf die Prognosefähigkeit seelenloser Algorithmen: Die ausführlichen Fragebögen der Online-Dating-Agenturen sollen die Wahrscheinlichkeit erhöhen, den passenden Partner zu finden.

# 37



**PROZENT**

**DER BEFRAGTEN DEUTSCHEN  
UNTERNEHMEN ERWARTEN DURCH  
OPTIMIERTE LOGISTIKPROZESSE  
GERINGERE KOSTEN.**

QUELLE: EXPERTON/BITKOM



**Datenexperte Saxe**

„Unser Ziel ist der intelligente Hafen“



Containerterminal im Hamburger Hafen: Eine komplexe Maschine, die sich im Idealfall selbst steuert

ROLAND MAGUNIA / DAPD

In Firmen, Forschungslabors und manchen Behörden setzt eine Art Goldrausch ein. Vielerorts wird das Mantra von den Daten als dem neuen „Öl“ oder „Gold“ des 21. Jahrhunderts besungen. Und schon jetzt profitieren einige von der Aufbruchstimmung: Statistiker, Physiker und sogenannte Data-Scientists oder Data-Miner, die Unternehmen in Sachen Big Data beraten. Ähnlich wie beim klassischen amerikanischen Goldrausch des 19. Jahrhunderts machen vor allem diejenigen gute Geschäfte, die Ausrüstung, Werkzeug und Expertise verkaufen, Big-Data-Spezialisten wie die 85-Mann-Firma Blue Yonder zum Beispiel.

## Wie Daten die Wirtschaft revolutionieren

Der Mann sieht nicht aus, als läse er auf Jahrmärkten aus Kristallkugeln, und doch sagt er mehrfach diesen seltsamen Satz: „Unser Job sind Vorhersagen aller Art.“ Uwe Weiss ist Geschäftsführer von Blue Yonder, einer Firma, die erst fünf Jahre alt ist. Weiss schaut nicht in Tarotkarten oder die Eingeweide unbescholtener Nutztiere, sondern auf die Zahlenkolonnen von Supermarktkassen, Wetterdiensten, Ferienterminen oder Verkehrsberichten. All das fließt in eine selbstentwickelte Datenanalyse-Software, die Blue-Yonder-Kunden wie die Otto-Gruppe oder die Drogeriemarktkette dm dann laut Eigenwerbung mit „präzisen Prognosen“ versorgt – etwa über den zu erwartenden Absatz eines bestimmten Artikels.

Für den Handel ist das extrem wichtig, denn so können Lieferschwierigkeiten oder Lagerkosten vermieden werden. Die

Software von Blue Yonder ist so programmiert, dass sie mit jeder eingehenden Information dazulernt und eigenständig Zusammenhänge erkennt.

So haben Weiss und seine Kollegen entdeckt, dass in einer Supermarktfiliale an manchen Tagen der Verkauf von Milch, Schoko-Riegeln oder Granny-Smith-Äpfeln in die Höhe schnellte – immer dann nämlich, wenn in der nahe gelegenen Jugendherberge neue Schulklassen ankamen. Die Software errechnet nun unter anderem anhand der Schulferien in den umliegenden Bundesländern, wie hoch die Wahrscheinlichkeit ist, dass neue Busladungen eintreffen.

Einen ähnlichen Aha-Effekt erlebten Blue-Yonder-Mitarbeiter bei einem Basisprodukt: aufgeschnittenem Brot. „An Brückentagen gehen Kinder nicht in die Schule, und schon wird weniger geschnittenes Brot gebraucht“, erklärt Weiss. Die Warenbestellsysteme passten sich vollautomatisch an. „Eine vergleichsweise einfache Sache.“

Durch den dauernden Dateneingang entsteht bei Blue Yonder so etwas wie eine Ad-hoc-Marktforschung über das Kaufverhalten, die auch für andere Zwecke eingesetzt werden kann. Die Drogeriekette dm lässt von der Weiss-Truppe ausrechnen, wie viel Personal sie in welcher Filiale einsetzen muss, auch den Umsatz pro Niederlassung prognostiziert Blue Yonder.

Für Versicherungen sind die Datenanalysen ähnlich interessant. So schildert Weiss als „Zukunftsszenario“ das mit mehr als 1000 Sensoren ausgestattete Auto, das unser Fahrverhalten permanent überwacht. Wer diese Daten, aus denen

sich leicht ein Risikoprofil erstellen lässt, seiner Versicherung zur Verfügung stelle, könnte künftig mit besonders günstigen Prämien gelockt werden, so Weiss.

„Big Data krepelt gerade die komplette Wirtschaft um, und wir stehen erst am Anfang“, sagt der Blue-Yonder-Chef, der das natürlich von Berufs wegen eher positiv sieht. Wichtige Big-Data-Kunden bestätigen aber erste messbare Erfolge.

IBM hat einige „Erfolgsgeschichten“ seiner eigenen Kunden in einer Studie zusammengefasst und berichtet darin von Effizienzsteigerungen um 20 Prozent. Beim Blue-Yonder-Kunden Otto heißt es, durch die Arbeit der Datenspezialisten habe sich „die Qualität der Absatzprognose für einzelne Artikel um 20 bis 40 Prozent“ erhöht. Die Versandhändler scheinen wirklich überzeugt zu sein, denn sie setzen das Verfahren inzwischen auch bei Unternehmensmarken wie Sportscheck ein – und haben sich zu 50 Prozent an Blue Yonder beteiligt.

Wie weit die Möglichkeiten reichen, zeigt auch der amerikanische Anbieter Netflix, der als DVD-Verleih begann und seine 36 Millionen Kunden mittlerweile hauptsächlich per Streaming mit Filmen versorgt. Mit der Serie „House of Cards“, einem Polit-Thriller mit Kevin Spacey in der Hauptrolle, feierte Netflix kürzlich neue Einschaltrekorde. Das war wohl kein Zufall, denn der Einkauf der Serie wurde mittels Datenanalysen abgesichert. Netflix als großes Streaming-Portal verfügt dafür über ideale Voraussetzungen: Die Firma weiß tagesaktuell, welche Genres gut laufen, wann Zuschauer das Interesse verlieren oder welche Schauspieler sie besonders mögen. So entsprach „House of

Cards“ exakt dem prognostizierten Publikumsgeschmack – die Rechnung ging auf.

Aus ähnlichen Gründen ist das Musik-Streaming-Portal Spotify beliebt. Die beteiligten Plattenfirmen bekommen aktuelle Einblicke in Musikgeschmack und Nutzungsverhalten, Bands können ihre nächste Tournee besser planen – indem sie an Orten gastieren, wo ihre Titel besonders oft abgerufen werden. Big Data verspricht aber noch ganz andere Hilfen.

### Ein Elektronenhirn soll den Krebs besiegen

Die Hoffnung für Millionen Krebspatienten wartet im zweiten Stock eines Forschungsinstituts. Es ist ein Gestell mit 25 Einschüben, an denen Dioden blitzen; die Rechner haben nicht wie üblich einen, sondern jeweils 40 Prozessoren. Der Raum wird eigens heruntergekühlt, damit das 1,5 Millionen Euro teure Gehirn mit seinen 1000 Prozessorkernen sich nicht heißrechnet.

**BIS ZU 40**  **PROZENT BESSERE VERKAUFSPROGNOSEN**

**ERREICHTE DIE OTTO-GRUPPE FÜR EINZELNE PRODUKTE DURCH DIE ARBEIT VON DATENSPEZIALISTEN.**



**Blue-Yonder-Chief Weiss**  
„Unser Job sind Vorhersagen aller Art“

Die Idee für das neuartige System ist hier in Potsdam-Babelsberg entstanden, das bislang vor allem für seine Filmstudios bekannt ist. Das Server-Gestell steht im Hauptsitz des Hasso-Plattner-Instituts (HPI), und tatsächlich war der SAP-Mitgründer, Sponsor und Namensgeber der Einrichtung selbst an der Entwicklung beteiligt. Die Idee wirkt simpel. Die Babelsberger Rechenmaschine saugt alle Daten direkt in ihren Arbeitsspeicher – und rechnet damit um den Faktor 1000 schneller als herkömmliche Geräte, bisweilen sogar noch mehr.

Das Verfahren, das hier als Projekt von acht Bachelor-Studenten mit dem Arbeitstitel „Sanssouci DB“ begann, ist international unter dem Namen „In-Memory“ bekannt. Es hat Innovationspreise gewonnen und es bis ins Portfolio von SAP geschafft. Die aktuelle Hana-Datenbank-Technologie des Walldorfer Konzerns basiert darauf. Der Chef des HPI, der Mathematiker Christoph Meinel, sieht neben den kommerziellen Möglichkeiten auch Chancen für die Krebstherapie. „Wir stehen dank des In-Memory-Verfahrens an der Schwelle zur personalisierten Medizin.“

Bislang dauere es Monate, das Genom eines Menschen zu entschlüsseln, um zu einer individuell auf den Patienten abgestimmten Therapie zu kommen – kein Wunder, bei rund drei Milliarden Erbgutbausteinen, die jeder Mensch in sich trägt. Heute weiß man allerdings, dass jeder Tumor verschieden ist und Therapien deshalb unterschiedlich wirken.

Mit Hilfe seines neuen „Superhirns“ lasse sich eine individuelle Genom-Bestimmung auf wenige Sekunden reduzieren, so Meinel. Zudem saugt der Babelsberger Computer über Nacht alle Informationen aus öffentlich zugänglichen Genom-Datenbanken und sucht darin nach vergleichbaren Fällen, um so Therapien ausfindig zu machen, die anderswo eine hohe Überlebensrate bei höchstmöglicher Lebensqualität sicherten. „Dieser Abgleich hätte bis vor kurzem noch Monate gedauert“, so Meinel.

Forscher der Universität Manchester arbeiten derzeit an einem anderen Big-Data-Projekt, einem „magischen Teppich“, der allein lebenden Senioren helfen könnte. Auf dem Fußboden wie ein normaler Teppich ausgelegt, registrieren die eingebauten Sensoren Schritte. So kann ein Rechner auswerten, ob die Person aufgestanden ist, und die Uhrzeit mit sonst üblichen Bewegungsmustern abgleichen. Wenn es Abweichungen vom üblichen Tagesrhythmus gibt, also ein medizinischer Notfall vorliegen könnte, wird Alarm geschlagen.

Die Teppich-Sensoren können kleinste Gang-Auffälligkeiten registrieren und so bereits vor einem möglichen Sturz anzeigen, dass etwas bei dem Patienten nicht stimmt. Auch mit einer logischen Weiter-

**75** 

**PROZENT**  
**DER DEUTSCHEN HABEN VON**  
**BIG DATA**  
**NOCH NIE GEHÖRT**  
**ODER KENNEN DIE BEDEUTUNG NICHT.**

QUELLE: BITKOM, ARIS

entwicklung wird bereits experimentiert: Denn natürlich können Chips und Sensoren mit derlei Alarmfunktionen auch direkt in künstliche Hüftgelenke eingebaut werden, um frühzeitig ein Warnsignal zu geben.

### Der Algorithmenbastler

Eigentlich wollte Stefan Henß sich mit Sportwetten etwas zum Studium hinzuverdienen. Er schrieb ein Programm, das Fußballergebnisse präzise vorhersagen sollte, um die Buchmacher auszutricksen. Das klappte eher leidlich.

Mit 10 hatte Henß seinen ersten Computer bekommen, mit 13 angefangen zu programmieren. Nun, vor drei Jahren, mit Anfang 20 und als Student an der TU Darmstadt, stieß er auf „Kaggle“ – eine Plattform, auf der Unternehmen Wettbewerbe ausschreiben. Es geht auch hier um möglichst konkrete Vorhersagen und knifflige Aufgaben, für die sie keine eigene Lösung finden.

Henß blieb an der Aufgabe einer Autohandelsplattform hängen, die nach einem Weg suchte, die Weiterverkaufsmöglichkeiten von Gebrauchtwagen zu prognostizieren. Er baute einen Algorithmus, in dem er zig Einzelinformationen über die Autos „in einen sinnvollen Kontext“ setzte, wie er es beschreibt, also Daten wie Erstzulassung, Kilometerstand, Jahreslaufleistung.

Der Hesse schickte seine Lösung ein und belegte damit den sechsten Platz unter 571 internationalen Teams, die um die



Otto-Verteilzentrum in Haldensleben

ausgesetzten 10 000 Dollar Prämie konkurrierten. Sein Ehrgeiz war geweckt: „Es war unglaublich motivierend, sich mit anderen zu messen, da habe ich Lunte gerochen“, so Henß, der seither zu einem der erfolgreichsten Algorithmenbastler bei Kaggle avancierte. Er sucht sich die Aufgaben inzwischen strategisch aus – auch nach Höhe des Preisgelds.

So kam er auch zu seinem bisher größten Triumph. Es ging um ein Programm, das automatisiert und zuverlässig die Aufsätze von Schülern bewerten sollte. Eine Benotungsmaschine sozusagen.

Henß baute aus diversen Standardalgorithmen ein Programm, das den Reichtum der Sprache berücksichtigt, den Anteil der Rechtschreibfehler pro Wort eruiert und auch Grammatikfehler erkennt. Es kann Rückschlüsse auf die Inhalte der Aufsätze ziehen. Sogar ob sie eher nüchtern geschrieben wurden oder Emotionen im Spiel waren, soll sein Algorithmus erkennen.

Eineinhalb Monate lang habe er an nichts anderem gearbeitet, am Ende bestand das Programm aus 12 000 Zeilen Code. Eine Woche vor Schluss des Wettbewerbs tat er sich mit zwei anderen Bewerbern zusammen, um die Gewinnchancen zu erhöhen. Die neuen Partner, neben Henß ein Engländer und ein Amerikaner, kannten sich nur über die Plattform und warfen ihre Lösungen zusammen. Sie gewannen. Das Preisgeld in Höhe von 60 000 Dollar teilten sie kollegial.

„Tests haben ergeben, dass sich unsere Bewertungen nicht signifikant von den Lehrerbewertungen unterscheiden“, sagt Henß. Mittlerweile hat das Trio die entstandene Software an das US-Unternehmen Pacific Metrics verkauft, Henß schreibt an seiner Master-Arbeit und muss sich um seine berufliche Zukunft wohl nicht sorgen. Doch es gibt auch Menschen, denen Big-Data-Anwendungen das Leben schwerer machen.

## Das Ende von Kommissar Zufall

Dümmer hätte es für den Autodieb kaum laufen können. Als er sich in einer Tiefgarage der kalifornischen Stadt Santa Cruz an einem Wagen zu schaffen machte, wurde er von einem Polizisten beobachtet, der wenige Meter weiter in einem zivilen Einsatzfahrzeug gerade an seinem Mittagessen kaute. Noch bevor der Wagen aufgebrochen war, wurde der Dieb festgenommen.

Der Streifenbeamte war nicht zufällig zur rechten Zeit am rechten Ort. Er verbrachte seine Mittagspause an diesem Tag in der Tiefgarage, weil ein Computerprogramm es ihm empfohlen hatte.

Seit rund zwei Jahren werden alle knapp hundert Polizisten der Stadt jeden Tag zu Dienstbeginn nicht nur von ihren Vorgesetzten, sondern von einem Algorithmus

# 40 000 000 000 000



## TERABYTE AN DATEN

**WERDEN IM JAHR 2020 WELTWEIT  
PRODUZIERT. IM JAHR 2012 WAREN ES  
2,8 MILLIARDEN TERABYTE.**

QUELLE: IDC

Die beiden Datenexperten Mohler und Brantingham haben mittlerweile eine Firma gegründet und vermarkten Predictive Policing weltweit. Allein in den USA haben bereits mehr als ein Dutzend Police Departments die Software eingeführt, darunter Los Angeles, Boston und Chicago. Jüngst war Clark in England, um der Polizei von Kent bei der Einführung zu helfen.

Und auch Militärs und Geheimdienste nutzen die Macht der Datenanalyse (siehe Seite 50). So spielte Big Data sogar bei der Jagd nach Osama Bin Laden eine wesentliche Rolle, wie der amerikanische Autor Mark Bowden in seinem Buch („The Finish“) beschreibt. Danach waren es unter anderem Datenbankanalysen, die die Fahnder schließlich ins pakistanische Abbottabad führten.

Einer der Ausrüster des amerikanischen Verteidigungsministeriums ist die 2004 gegründete Firma Palantir, hinter der unter anderen der deutsch-amerikanische Investor Peter Thiel steckt. Unter Geheimdienstlern und Militärs ebenfalls heiß gehandelt wird die kalifornische Software-Schmiede Splunk, deren Zentrale in einer ehemaligen Wurstfabrik in San Francisco untergebracht ist. Vor wenigen Wochen wurde Splunk von Technologiejournalisten unter die fünf innovativsten Unternehmen der Welt gewählt – Google rangierte nur auf Platz elf.

Ihre Anwendungen nutzen Regierungen, Behörden und Unternehmen aus über 90 Ländern. Auf dem amerikanischen Heimatmarkt gehören zu den Splunk-Kunden auch das Pentagon und die Heimatschutzbehörde. Die Software des neun Jahre alten Unternehmens analysiert und interpretiert Daten von Maschinen – egal welcher Art: Mobilfunktürme, Klimaanlage, Web-Server oder Flugzeuge.

„Die Turbinen eines Airbus A380 produzieren während eines Flugs so viele Daten wie ein mittelgroßes Rechenzentrum“, sagt Guido Schröder, der Produktchef von Splunk.

Schröder kommt ursprünglich aus Paderborn, war lange bei SAP, nun hat er 160 Entwickler und Ingenieure unter sich, deren Ziel es ist, die von Maschinen produzierte Datenflut für Behörden, aber auch für Unternehmen nutzbar zu machen.

auf ihre Schicht vorbereitet. Ständig neu gefüttert mit allen relevanten Daten, die der Polizeiapparat zu bieten hat, errechnet das Programm, in welchen Gegenden der Stadt zu welcher Uhrzeit die höchste Wahrscheinlichkeit für eine Straftat besteht – etwa Einbruch, Raub oder Autodiebstahl. Tötungsdelikte sind bislang ausgeschlossen.

Die 15 gerade gefährlichsten Gebiete tauchen als Quadrate auf elektronischen

Karten auf. In zwei Dritteln der Fälle gab es bisher tatsächlich verdächtige Vorkommnisse. „Ich wäre schon mit zehn Prozent glücklich gewesen“, sagt der stellvertretende Polizeichef von Santa Cruz, Steve Clark.

Die vorhersehende Kriminalitätsbekämpfung wurde maßgeblich von zwei Professoren entwickelt, dem Computerwissenschaftler George Mohler und dem auf Verbrechen szenarien spezialisierten Anthropologen Jeffrey Brantingham. Das Programm der beiden basiert auf Modellen zur Vorhersage von Nachbeben nach Erdbeben.

Steve Clark hatte Anfang 2011 zufällig von der Idee der beiden Wissenschaftler gehört. Gemeinsam setzten sie ein Pilotprojekt auf. Sie fütterten das Programm mit Verbrechenstatistiken aus acht Jahren und zahllosen anderen Daten, die eine Rolle spielen können: das Wetter, die Nähe zu Parks und Buslinien. Dazu muss jedes Verbrechen mit jedem anderen in Beziehung gesetzt werden.

„Es gab anfangs viele Skeptiker, mich eingeschlossen“, sagt Clark. „Aber die Zahlen sprechen eine klare Sprache: Es funktioniert.“ Nach dem ersten Jahr mit der neuen Verbrechenvorhersage sei in Santa Cruz die Zahl der Einbrüche um elf Prozent und die der Autodiebstähle um acht Prozent zurückgegangen. Auch die Zahl der Festnahmen hat sich in Santa Cruz deutlich erhöht – um 56 Prozent.

Inzwischen ist das ganze Polizeipräsidium technisch hochgerüstet. Alle Cops sind mit Smartphones und Tablets ausgestattet, um von unterwegs auf das webbasierte Vorhersageprogramm zugreifen zu können. Sie sind angehalten, wann immer es geht, Zeit in den markierten Gebieten zu verbringen. Clark hat reihenweise Geschichten parat, wie Einbrecher und Diebe von seinen Leuten in flagranti in einer der vorhergesagten Zonen erwischt wurden. Er lacht, wenn er die Erfolgsgeschichten des Kollegen Computer erzählt.

# 4,4



## MILLIONEN NEUE JOBS

**IM IT-BEREICH DURCH BIG DATA,  
WELTWEIT, BIS 2015**

QUELLE: GARTNER



NOBERT VON DER GROEBEN / DER SPIEGEL

**Santa-Cruz-Cop Clark (r.), Kollege**  
„Die Zahlen sprechen eine klare Sprache“

Im Fall des A380 kann es darum gehen, den Kerosinverbrauch zu minimieren oder die Wartungsintervalle zu optimieren.

„Sicherheit ist eines der größten Wachstumsgebiete für Big-Data-Anwendungen“, sagt Schröder. Splunk hat dabei nicht nur Kriminalität und Terrorismus im Blick, sondern auch die zunehmenden Attacken im und über das Internet.

So kann die Software beispielsweise Hackerangriffe erkennen oder andere Cyber-Attacken. „Wir bringen uns in Stellung für einen ausufernden Cyberwar.“ Aber der neue Krieg der Datenjäger hat auch vielfältige zivile Seiten.

## Die Daten-Bank

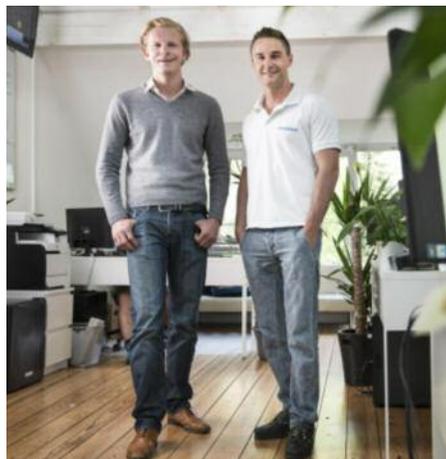
Der Backsteinbau in Hamburg-Winterhude sieht nicht aus wie eine Bankfiliale, und doch werden hier Kredite vermittelt. „Kreditech“ steht auf einem Plastikbanner, drinnen herrscht eine Atmosphären-Mischung aus Garagenfirma und Wohn-gemeinschaft, überall hängen Zettel an den Wänden: „Schuhe aus“.

Sebastian Diemer und Alexander Graubner-Müller treten nicht wie typische Vertreter deutscher Kreditinstitute auf, deren Geschäftsmodell sie ohnehin für gestrig halten. „Zumindest das klassische Filialbankmodell ist ein Auslaufmodell“, sagt der jung-dynamische Diemer, der den klassischen Start-up-Look trägt: Jeans und Kapuzenpulli.

Die selbstbewussten Firmengründer verleihen Geld übers Internet, kurzfristige Mikrokredite bis maximal 500 Euro, wobei der Durchschnittskunde 109 Euro erhält. Sie verlangen dafür keine Schufa-Auskunft, sondern ermitteln die Ausfallwahrscheinlichkeit selbst – ihr Social-Scoring-Verfahren besteht aus superschnellen Datenanalysen. „Im Idealfall ist das Geld innerhalb von 15 Minuten nach Bewilligung auf dem Konto, in Polen klappt das schon regelmäßig“, so Diemer. Kreditech will dafür von seinen Nutzern so viele Daten wie irgend möglich – je mehr das Unternehmen bekommt, desto präziser wird die Prognose, desto höher der mögliche Kreditrahmen für die Kunden.

Die Bewertungsprofile von Ebay-Konten sind ohnehin öffentlich einsehbar. Zudem verlangt Kreditech Zugang zu Facebook-Profilen, um zu prüfen, ob Foto und Ort mit anderen Informationen etwa bei Xing oder LinkedIn übereinstimmen – und ob es unter den Freunden viele mit ähnlichen Bildungsabschlüssen gibt oder viele Kollegen, die in derselben Firma arbeiten.

Das alles erhöhe die Wahrscheinlichkeit, es mit einer realen Person zu tun zu haben, erklärt Diemer. Selbst die Frage, ob die Kreditbitte von einem teuren iPad aus gestellt wurde oder von einem billigen Aldi-Computer, geht mit in die Wertung ein. Auch das Verhalten der Antragsteller spielt eine Rolle – etwa die Zeit,



Kreditech-Gründer Graubner-Müller, Diemer  
„In Polen klappt das schon“



**MILLIARDEN  
KREDITKARTEN-  
TRANSAKTIONEN  
WURDEN 2012 GETÄTIGT.**

QUELLE: NILSON REPORT

die sie zum Ausfüllen des Fragebogens brauchen. Ebenso werden die Fehlerhäufigkeit und der Einsatz der Löschtaste nüchtern registriert.

Bis zu 8000 verschiedene Einzelinformationen verarbeite der Kreditech-Algorithmus auf diese Weise pro Anfrage, sagen die Macher, die ihre Firma im März 2012 gründeten und schnell expandieren. In Polen, Spanien, Tschechien ist Kreditech schon online, in Russland wollen die Hamburger in diesen Tagen starten.

In Deutschland lief der Geldverleih nur drei Wochen – dann stellte Kreditech den Service wieder ein. „Präventiv“, wie Diemer sagt. Tatsächlich hatte sich da bereits die Finanzaufsicht BaFin gemeldet und eine Prüfung des Geschäftsmodells angekündigt. Die Hamburger verlangten näm-

lich nicht nur hohe Zinsen, sondern auch den Erwerb eines „Bonitätszertifikats“, das bis zu 49 Euro kosten konnte.

Den Zwangserwerb von Zertifikaten habe man inzwischen auch in den anderen Märkten abgeschafft, so die Kreditech-Macher. Die von ihnen verlangten Zinsen rangierten je nach Darlehenshöhe, Score-Wert und Laufzeit zwischen 5 und 28 Prozent im Monat.

Es sind indes nicht die Mikrokreditanfragen und Zinseinkünfte, von denen sich die Macher hohe Umsätze versprechen. Ihr eigentliches Ziel ist der Aufbau einer internationalen, sich selbst aktualisierenden Bonitätsdatenbank für andere Unternehmen, Online-Händler beispielsweise. Bisherige Scoring-Verfahren basierten auf weniger Parametern, die zudem nur die Kreditvergangenheit abbildeten, so Diemer. In vielen Ländern gebe es nicht mal das – eine Schufa mit einer ähnlich großen Marktdurchdringung wie in Deutschland existiert nur in wenigen Märkten: „Für fast drei Viertel der Weltbevölkerung gibt es noch keine verlässlichen Bonitätsauskünfte“, so Graubner-Müller.

Neben Kreditech versuchen sich auch Anbieter wie Zestfinance und das britische Wonga mit ähnlichen Zielen in dem prekären Markt, der neben rechtlichen Fragen auch moralische aufwirft. In Großbritannien machte Wonga Negativschlagzeilen, als es versuchte, Studenten aus staatlichen Studentenkrediten in eigene, ungleich teurere Darlehen zu locken.

Die Kreditech-Gründer, die sich schon seit ihren Jugendtagen in Wiesbaden kennen, beteuern, man habe sich weder in Sachen Datenschutz noch bei der Zinshöhe etwas vorzuwerfen. „Die Schufa speichert Daten auf Vorrat, wir nur bei einer konkreten Anfrage.“ Im Übrigen würden die Daten abgelehnter Antragsteller nach 90 Tagen fast sämtlich gelöscht – nur was notwendig sei, um einen einmal abgelehnten Bewerber wiederzuerkennen, werde länger vorgehalten.

Investoren halten Social Scoring offenbar dennoch für außerordentlich attraktiv: Kreditech sicherte sich im Dezember vier Millionen Dollar, erst im April stiegen zudem die Samwer-Brüder mit einem Fonds in ähnlicher Höhe ein. Wonga sammelte bereits 141 Millionen Dollar von Geldgebern ein.

## Eine Tyrannei der Daten?

Geschäftsmodelle wie das von Kreditech zeigen exemplarisch die Brisanz der Fragen, die all die neuen Big-Data-Anwendungen aufwerfen. Natürlich geben Nutzer ihre Daten Schritt für Schritt „freiwillig“ preis – so wie wir bei Facebook freiwillig und teils freizügig Privatfotos posten oder über Twitter unsere politischen Ansichten verbreiten. Jeder ist letztlich ein Lieferant des großen, neuen

Datenrohstoffs, auch in der analogen Welt, wo wir Payback-Karten benutzen, Meilen sammeln, Autos mieten.

Vielleicht tun das auch deshalb viele so unbefangen, weil meist intransparent bleibt, was mit unseren Daten passiert. An wen und wie oft werden sie weiterverkauft? Gelten dort auch Lösch- oder Anonymisierungsregeln? Und was passiert beispielsweise mit den Bonitätsprofilen von Kreditech, sollte die kleine Firma mal von einer größeren übernommen werden oder pleitegehen?

Wie sensibel die Öffentlichkeit inzwischen auf derlei Fragen reagiert, zeigte der Versuch der etablierten Bonitätsprüfer der Schufa, gemeinsam mit dem Hasso-Plattner-Institut ein Pilotprojekt zum Social Scoring zu starten.

Es ging wie bei Kreditech darum, Daten aus Facebook, Twitter und anderen sozialen Netzwerken auszuwerten und deren Aussagekraft hinsichtlich der Kreditwürdigkeit zu erforschen. Schon die Ankündigung des Projekts löste gewaltige Proteste aus. Das Vorhaben wurde schleunigst beerdigt.

Noch größer war der Sturm der Entzündung, als vielen Autofahrern klar wurde, dass ihr Navigationsgerät nicht nur den schnellsten Weg zum Ziel kennt, sondern auch gegen sie eingesetzt werden kann. Der niederländische Navi-Hersteller TomTom hatte seine Daten an die niederländische Regierung verkauft. Die gab das Material an die Polizei weiter, die es nutzte, um ihre Radarfallen möglichst gewinnbringend aufzustellen – an Orten, an denen besonders viele TomTom-Nutzer zu schnell unterwegs gewesen waren.

Der TomTom-Chef entschuldigte sich öffentlich. Man habe geglaubt, der Regierung sei an mehr Verkehrssicherheit und Stauvermeidung gelegen gewesen. Die Radarfallen-Anwendung habe TomTom „nicht vorhergesehen“.

Die nächsten derartigen Kollisionen sind programmiert, zumal der Entwicklung ein zentraler Konflikt innewohnt. Besonders wertvoll sind Big-Data-Anwendungen, wenn sie personalisiert sind – wie die Bonitätsprüfung oder die individuelle Medizin.

Personalisierte Profile, die eine Fülle von Informationen zusammenfassen, von der Meinungsäußerung bei Facebook bis zum Bewegungsprofil, eröffnen Firmen verlockende Möglichkeiten. Wer bei Facebook eine Jeans „geliked“ hat, dem könnte der Ladeninhaber beispielsweise beim Eintreten einen Rabattcoupon für genau dieses Modell aufs Handy schicken.

Aus Sicht von Händlern und manchen Verbrauchern mag das attraktiv sein, Datenschützer bewerten viele Big-Data-Ideen dagegen als Big-Brother-Szenarien in einer völlig neuen Dimension.

Bislang bemühen sich viele Unternehmen, derlei Bedenken mit dem Hinweis

4,6 

**MILLIARDEN  
EURO UMSATZ  
WURDE WELTWEIT IM JAHR  
2012 MIT BIG-DATA-  
ANWENDUNGEN GEMACHT.  
BIS 2016 SOLL ER AUF KNAPP  
16 MILLIARDEN  
ANSTEIGEN.**

QUELLE: BITKOM

zu zerstreuen, die Daten nur „anonymisiert“ zu erheben, zu speichern und zu analysieren – was allerdings die Macht der Datenanalysen offenbar gehörig unterschätzt.

Die Vorstellung von anonymisierten Bewegungsprofilen zum Beispiel sind wohl ein Mythos: Nach einer aktuellen Untersuchung des „Scientific Reports“ sind unsere Mobilitätsmuster so verschieden, dass sie recht einfach und mit einer Trefferquote von 95 Prozent Individuen zuzuordnen sind. Je mehr Daten kursieren und analysiert werden können, desto wahrscheinlicher wird, dass Anonymität „algorithmisch unmöglich“ wird, glaubt der Princeton-Wissenschaftler Arvind Narayanan.

In seinem Blog beschreibt er, dass schon 33 Bits an Information ausreichen, um eine Person zu identifizieren.

Aus Sicht von Firmen ist die leicht schizophrene Haltung von Verbrauchern die eigentliche Crux: Einerseits sind sie online – und öffentlich einsehbar – unheimlich mitteilbar, gleichzeitig hegen sie düsterste Datenschutzbedenken.

So spricht eine Studie für die New Yorker Werbeagentur Ogilvy One von „unsexy Datenängsten“. 75 Prozent aller Befragten wollen danach keine persönlichen Daten bei Unternehmen gespeichert wis-

sen. Fast 90 Prozent sprachen sich dagegen aus, dass Firmen ihre Surf-Bewegungen im Internet „tracken“, also überwachen.

Dieser Konflikt erklärt die Schärfe, mit der gerade um die neue europäische Datenschutzverordnung gestritten wird. Werden die Pläne der EU-Kommission Realität, die auch ein „Recht auf Vergessenwerden“ im Netz beinhalten, sehen viele Anbieter ihre Big-Data-Wachstumsphantasien in Gefahr.

Auch deshalb erlebt Brüssel gerade eine Lobbykämpfe von Amazon, Google, Facebook und Co.

Noch drängender allerdings muss eine moderne Gesellschaft sich fragen, ob sie alles akzeptieren soll und will, was in der datengetriebenen Wirtschaft möglich wird. Wollen wir in einer Welt leben, in der Algorithmen vorhersagen, wie gut ein Kind in der Schule sein wird, wie geeignet für einen bestimmten Job – oder wie gefährdet es ist, straffällig zu werden oder an Krebs zu erkranken?

Ist es wirklich erstrebenswert, dass Kulturgüter wie TV-Serien oder Musikalben mittels datengetriebener Analysen nach unseren prognostizierten Wünschen maßgeschneidert werden? Wo bleibt in dieser total vermessenen und berechneten Welt die Kreativität, die Intuition, das Überraschungsmoment?

Der Netz-Philosoph Jewgenij Morosow warnt vor einer drohenden „Tyrannei der Algorithmen“ und kritisiert ganz grundsätzlich die Ideologie, die hinter vielen aktuellen Big-Data-Anwendungen steht.

Weil zunehmend auch Formeln im Finanzbereich und wie bei Predictive Policing in der Polizeiarbeit eingesetzt würden, müssten sie von unabhängigen, qualifizierten Auditoren regelmäßig geprüft werden – schon um Diskriminierung und Machtmissbrauch zu verhindern.

Wie überfällig eine breite gesellschaftliche und politische Debatte darüber ist, ließ einer der dominierenden Big-Data-Giganten einmal eher beiläufig durchblicken.

Man habe mit dem Gedanken gespielt, mittels der einlaufenden Suchanfragen Börsenkurse vorherzusagen, erzählte der heutige Verwaltungsratsvorsitzende von Google, Eric Schmidt, schon im Jahr 2010. Die Idee habe man allerdings wieder verworfen, weil man intern zu der Auffassung kam, dass das wohl illegal sei.

Schmidt sagte, dass man es eben lasse. Er sagte nicht, dass es unmöglich sei.

MARTIN U. MÜLLER, MARCEL ROSENBACH,  
THOMAS SCHULZ



**Video: So funktioniert die Verbrechensthematik der Polizei**

spiegel.de/app202013daten  
oder in der App DER SPIEGEL