

Der Abhörskandal um die US-Spionagebehörde NSA hat es offengelegt: Wenn man sie zu entschlüsseln weiss, steckt in den Datenbergen, die wir täglich produzieren, viel Wissen, Geld und Macht. Doch Big Data ist nicht nur eine Gefahr. Sondern auch eine Chance, unser Leben zu verbessern. SEITE 28 + 29

Das Erdöl der nächsten Revolution

Zahlen, Wörter, Bilder, Töne: 2014 werden wir doppelt so viele Daten produzieren wie 2012. Ob wir wollen oder nicht. Der unvorstellbar schnell wachsende digitale Datenberg ist ein Rohstoff, von dessen Abbau man sich Milliardenprofite verspricht – wenn es den Big-Data-Experten gelingt, Technologie und Ethik miteinander zu versöhnen.

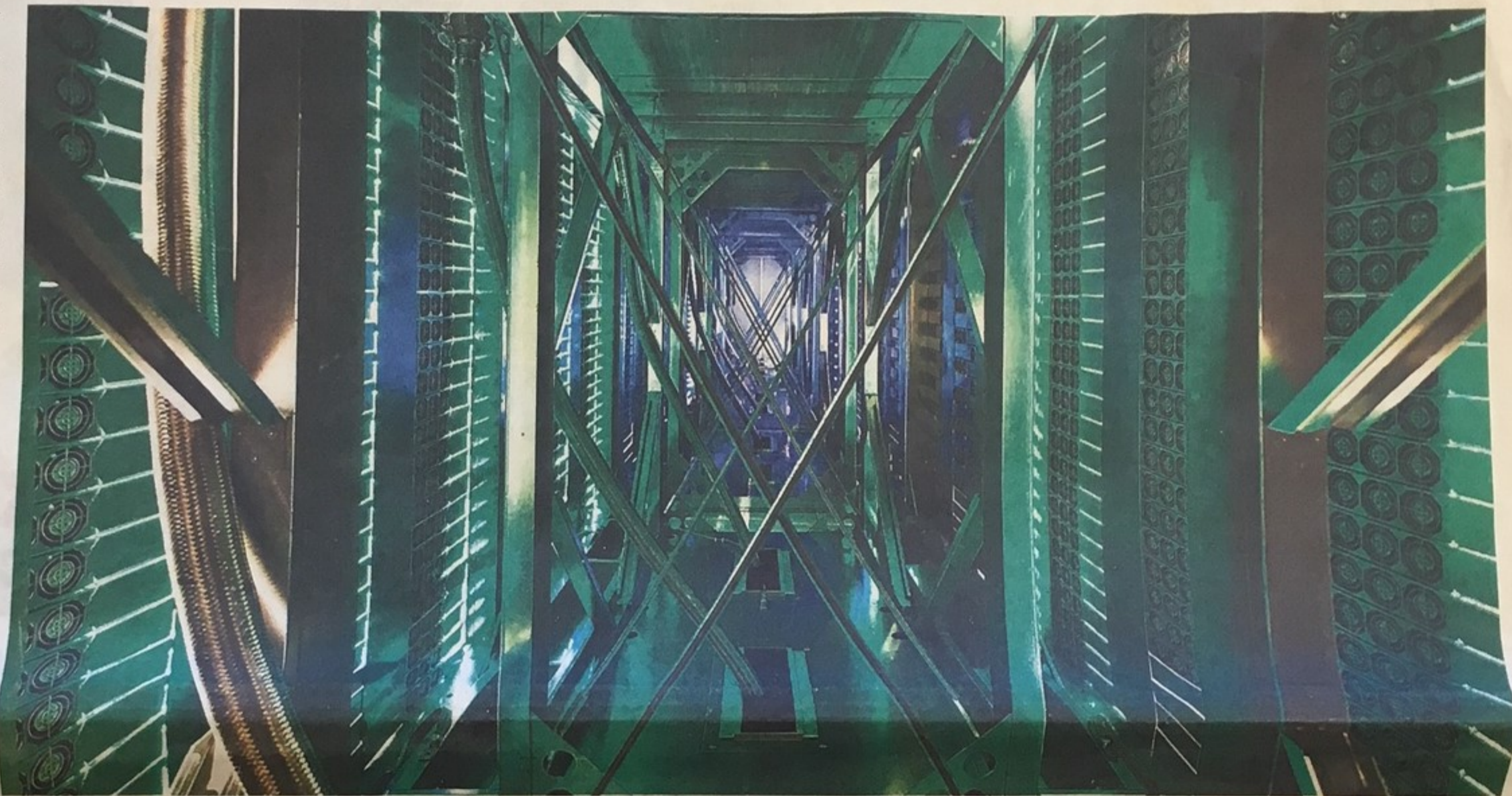
Man stellt sich unter Big-Data-Revolutionären gerne bleiche, weltabgewandte Nerds vor, die zwischen Bergen von Pizzaschachteln, eingerahmt von blinkenden Flachbildschirmen, mit flackerndem Blick Tastatur und Maus bearbeiten. Die mit keinem Menschen reden, schwarze Gothic-T-Shirts tragen, die Körperpflege vernachlässigen, aber plötzlich aus unübersichtlichen Datenkolonnen ein überraschendes Ergebnis fischen – für das sich die Überwachungsbehörden interessieren könnten.

Kurt Stockinger und Thilo Stadelmann sehen nicht aus wie Nerds. Sie sind elegant gekleidet, eloquent, geerdet, frisch geduscht – und nach dem Gespräch gehen sie ins Restaurant zum Mittagessen. Computerfreaks sind sie trotzdem, und Big Data war schon ihr Thema, bevor es den Begriff gab. Big Data, sagen sie cool, seien im Prinzip alle Datenmengen, die so gross seien, dass sie nicht mehr in eine Excel-Tabelle passten. Stockinger und Stadelmann arbeiten am Institut für angewandte Informationstechnologie der Zürcher Hochschule für angewandte Wissenschaften – oder neudeutsch: im ZHAW-Departement School of Engineering – in Winterthur.

Sie gehören zur Gruppe von Wissenschaftlern, die man als schweizerische Avantgarde der Grossdatenanalyse bezeichnen könnte. Und die den Ehrgeiz haben, mit der Weltspitze in dieser Zukunftsdisziplin im kalifornischen Silicon Valley mithalten zu wollen. Im Frühjahr 2013 gründete ein internationales Team um Stockinger und Stadelmann nach US-Vorbild an der ZHAW School of Engineering das Datalab, ein interdisziplinäres datenwissenschaftliches Labor, eines der ersten in Europa.

Um den aktuellen Datenbestand Westeuropas zu speichern, müsste jeder Einwohner 12 000 USB-Sticks mit einem Gigabyte Kapazität auf sich tragen. Pro Minute werden weltweit 40 Videostunden auf Youtube hochgeladen, 700 000 Nachrichten auf Facebook gepostet und 100 000 Twittermeldungen abgesetzt. So häufen wir alle täglich Big Data an. Das Winterthurer Datalab ist ein Ort, an dem man erfahren kann, was aus diesen riesigen Datenbergen herauszuholen ist. Und was sie aus uns herausholen.

Big Data geht nur mit neuen Köpfen. Das Datalab ist ein virtuelles Forschungslabor, an dem Informatiker, Mathematiker, Statistiker und Ingenieure zusammenarbeiten, «und diese fachlichen Grenzüberschreitungen», sagt Kurt Stockinger, «sind ein zentraler Aspekt des neuen Datenzeitalters». Eindimensionales Denken von Computernerds früherer Prägung sei den



Daten sind eine rasant nachwachsende Ressource des industriellen Fortschritts: So sieht ein Rechenzentrum des Internetriesen Google in den USA aus.

Keystone

heutigen riesigen Datenbergen nicht mehr gewachsen. Moderne Datenwissenschaftler müssten in Teams kooperieren, sich schnell in fremde Fachgebiete hineinendenken, unternehmerische Fragen mitbedenken. Vor allem aber auch gut kommunizieren können, sagt Stockinger. Die beste Datenauswertung nütze nichts, wenn man als Analytiker dem Firmenchef nicht erklären könne, für welche Entscheidungen die Ergebnisse nützlich seien. Und für welche nicht.

Die Nachfrage der Wirtschaft nach interdisziplinären Datenwissenschaftlern, wie sie die School of Engineering im Bachelorstudium Informatik und Wirtschaftsingenieurwesen sowie in entsprechenden Weiterbildungsprogrammen ausbilde, wachse derzeit «extrem stark», sagt Stockinger – obschon die Analyse riesiger Datenmengen eigentlich nichts Neues ist. In den Naturwissenschaften – etwa bei Modellanalysen zur Klimaerwärmung, am Teilchenbeschleuniger Cern bei Genf oder in der Genom-Forschung – wird schon seit Jahrzehnten datenintensiv gearbeitet. «Wirklich neu ist», sagt Thilo Stadelmann, «dass in den letzten paar Jahren vor allem in der Wirtschaft eine neue mentale Einstellung Fuss gefasst hat: Früher als unwichtig verworfene Daten werden nun inspiziert. Denn man hat erkannt und ist nun überzeugt davon, dass Entscheidungen besser werden, wenn man sich nicht nur aufs Bauchgefühl stützt. Sondern auf Daten.» Selbst wenn diese Daten nur zufällig anfallen.

Eines der grossen Pionierfelder von Big Data ist das Onlinemarketing. Um ihre Kunden besser kennen zu lernen, haben Marktführer des Internetgeschäfts – etwa die US-Riesen Amazon oder Google – begonnen, Informationen über uns, die wir auf verschiedenen Kanälen gewollt oder ungewollt hinterlassen, systematisch zu analysieren. Heute kann man davon ausgehen, dass ein Onlineanbieter viel über einen weiss, selbst wenn man sich das erste Mal bei ihm einloggt.

Im öffentlichen Bewusstsein werde Big Data zu oft limitiert darauf, wie man Konsumenten noch raffinierter verführen könne. «Aus diesem Stadium», sagt Kurt Stockinger, «hat sich die Datenwissenschaft inzwischen herausgearbeitet.» Salopp gesagt: Moderne Datenspezialisten verstehen sich nicht mehr als Vollzugsgehilfen globaler Marketingstrategien, sondern als eigenständige Treiber einer technologischen Revolution, die vor keinem Lebensbereich haltmacht. Big Data überwacht medizinische Operationen, Verkehrsflüsse, Energieversorgungsnetze. Big Data steuert Investitionsentscheide, CEO-Ernennungen, politische Wahlen. Big Data lenkt sportliche Spitzenleistungen, Burn-out-Präventionskampagnen, Ernteversicherungen (siehe Seite 28 und 29).

Genährt wird der Big-Data-Mythos durch eine für Laien unverständliche Sprache. Man sieht schon klarer, wenn man drei Stichwörter kennt: Hadoop, SAP, Algorithmus. Die riesigen

Datenmengen müssen auf verschiedene Server abgelagert und parallel bearbeitet werden. Hadoop ist ein Softwarepaket, das diesen Entwicklungsschritt möglich gemacht hat. Der deutsche Softwarekonzern SAP – den viele Angestellte vom Lohnabrechnungs- und Ferienguthaben-administrator SAP online kennen – hat mit seiner 2010 lancierten revolutionären Datenbanktechnologie Hana die Geschwindigkeit von Big-Data-Analysen enorm erhöht, sodass unüberblickbare Datenmengen in Sekundenschnelle durchforstet werden können.

Unter einem Algorithmus versteht man die Befehlskette, die der Computer braucht, um bestimmte Datenanalysen auszuführen. «Algorithmen funktionieren eigentlich wie Kochrezepte», sagt Kurt Stockinger, «sie legen Zutaten und Mengenverhältnisse fest, die man braucht, um Daten sinnvoll zueinander in Beziehung zu setzen.» Dass Datenwissenschaftler neue Algorithmen erfinden, komme zwar vor, sei aber eher selten. Das Expertenwissen bestehe häufig in der Erfahrung, welche existierenden Algorithmentypen sich für die Bearbeitung welcher Datentypen eignen.

Algorithmen erkennen in Informationsfluten Muster, die sich der menschlichen Auffassungsfähigkeit entziehen. Ob die gefundenen Zusammenhänge plausibel sind oder nicht, müssen dann aber wieder Menschen beurteilen. «Das ist die grosse Herausforderung», sagt Thilo Stadelmann. Moderne Datenwissen-

Moderne Datenspezialisten verstehen sich als Treiber einer technologischen Revolution.

schaftler müssten fähig sein, den Boden der Realwelt nicht unter den Füßen zu verlieren und rechtliche, ethische oder soziale Fragen im Auge zu behalten.

Schub gibt der entstehenden Big-Data-Industrie die Tatsache, dass Algorithmen es heute verstehen, Datenformen miteinander zu kombinieren, die bisher nicht verknüpft werden konnten – Geschäftszahlen mit Social-Media-Kommentaren zum Beispiel. Das befeuert auch die Gründung neuer, ambitionierter Firmen. Das Zürcher Start-up Squirro zum Beispiel versucht sich auf dem Markt zu etablieren mit einer Suchtechnologie, dank der Firmen unter anderem Auswirkungen von Twittermeldungen über eigene Produkte auf den Aktienkurs in Sekundenschnelle früherkennen – und allenfalls Gegensteuer geben können. Das Münchner Start-up Soma Analytics hat eine Anwendung entwickelt, mit der das Smartphone Stresslevel und Schlafqualität der Mitarbeiter überwacht und Unternehmungen rechtzeitig Burn-out-Präventionsprogramme starten können.

Die rasante Entwicklung, sagt Thilo Stadelmann, lasse den Wert von Daten unglaublich steigen – selbst wenn man noch gar nicht genau wisse, was man mit ihnen anfangen wolle. Firmen beginnen, Datenbesitz als Aktivposten in der Bilanz aufzuführen. Die Rechnung, die gemacht wird, ist laut Stadelmann die folgende: Brauchbare Hardware kostet vielleicht 5000 Franken. Software ist zehn- bis hundertmal so viel wert wie Hardware. Und Daten sind zehn- bis hundertmal so viel wert wie Software. Das macht sie zum Erdöl der Zukunft.

Und die Revolution hat – in den Augen der Datenspezialisten – gerade erst begonnen. Heute werden Big-Data-Analysen meistens von Menschen gesteuert, die Konsequenzen daraus von Menschen gezogen. Das muss nicht so bleiben. Im künftigen «Internet der Dinge» vernetzen sich intelligente Computer selber, ziehen aus den analysierten Daten ihre eigenen Schlüsse. Sensoren melden, wenn man ins Büro kommt, und werfen alle nötigen Geräte an. Zu Hause beobachten Sensoren, wie draussen das Wetter ist, und richten automatisch Stores, Heizung und Lüftung danach aus. Smarte Autos fahren automatisch durch den Verkehr, während der Lenker Zeitung liest oder E-Mails abarbeitet.

Könnte sein, dass wir nicht lange Zeit haben, uns darauf vorzubereiten.

Jürg Steiner

juerg.steiner@bernerzeitung.ch