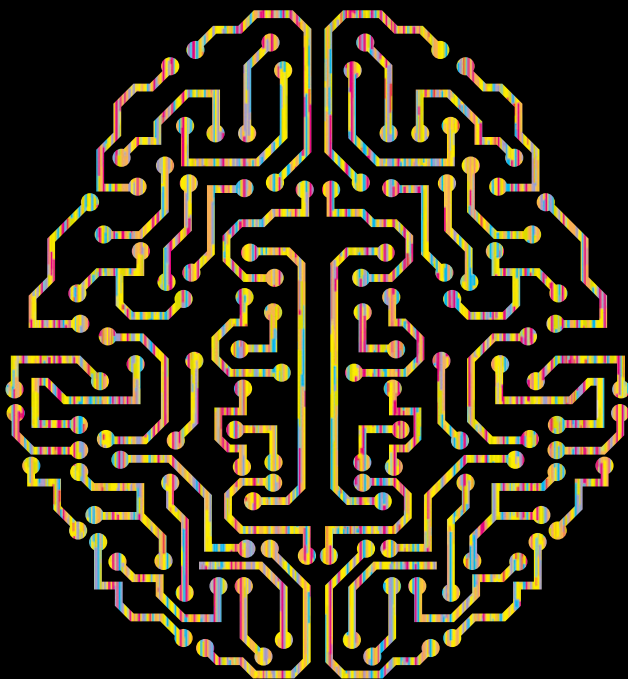


DIE DATENWELT

Zwischen Rausch und Klarheit



SZENARIEN FÜR DIE

DATENGESELLSCHAFT

Die Euphorie in Forschung und Wirtschaft ist riesig. Ein grosser Entwicklungsschritt ist im Gange: Die Maschinen beginnen zu denken. Jahrzehntlang wurde die Welt vermessen und Massen von Daten gespeichert, nun stehen Methoden zur Verknüpfung und Auswertung bereit. Algorithmen lassen die Informationen aus den vormals stummen Datenbergen sprudeln und beeinflussen unser

aller Leben – von den Resultaten bei der Google-Suche über die Entdeckung des Higgs-Teilchens bis hin zum Börsencrash. Doch wer genauer hinschaut, erkennt, dass wir das Denken nicht den Computern überlassen können: Je schlauer die Maschinen werden, desto mehr müssen wir auch selber mitdenken. Denn was Sinn macht, kann letztlich nur der Mensch entscheiden.

I. VERMESSEN

Die reale Welt wird in Daten übertragen. Im Jahr 2011 wurden weltweit über 2 Zettabyte neue Daten generiert. Studien schätzen das 2012 neu hinzugekommene Datenvolumen auf 2,7 Zettabyte. Laut McKinsey & Co. soll diese Menge bis 2020 jährlich um 40% ansteigen. Die Datenmenge wächst schneller als die Speicherkapazitäten.

WAS WIR DENKEN:

2,9 Mio. E-Mails werden pro Sekunde weltweit versandt, pro Minute 660 000 neue Einträge auf Facebook geschaltet.

WAS WIR FÜHLEN:

Rund 35 000 individuelle Likes pro Minute auf Facebook sowie unzählige Emoticons in Textnachrichten zeigen momentane Gefühlszustände.

WO WIR SIND:

GPS in Mobiltelefonen zeigt Bewegungsverläufe; pro Minute erfolgen über 21 000 Check-ins allein auf dem persönlichen Lokalisierungsdienst Foursquare.

- ! Zu Beginn des Jahres 2012 waren nur noch 2% aller Daten nicht digital gespeichert.
- ! 90% der heute weltweit vorhandenen Daten wurden erst in den letzten zwei Jahren generiert.
- ! 2050 wird jedes Individuum – bei einer geschätzten Weltbevölkerung von neun Mia. – durchschnittlich 489 Terabyte an persönlichen Daten erzeugt haben.

WAS WIR EINKAUFEN:

Händler und Zahlungsmittelanbieter wie Paypal oder Visa speichern alle Transaktionen. Apple zählt 47 000 App-Downloads pro Minute.

WAS WIR SEHEN:

Jede Minute werden auf Youtube 48 Stunden neue Videos hochgeladen und rund 7000 Bilder auf Flickr und Instagram.

WAS WIR SUCHEN:

Allein Google erhält pro Minute 2 Mio. Suchanfragen von seinen Nutzern. Dazu kommen Anfragen bei Baidu, Yahoo, Bing etc.

II. VERKNÜPFEN

Die Zusammenführung der allerorten erhobenen Datenmengen ist die Achillesferse des Daten-Hypes. Oft entstehen nützliche Informationen erst durch die Verknüpfung verschiedener Datensätze – zum Beispiel, wenn die Kombination von Orts- und Geschwindigkeitsangaben von Autos anzeigen, wo gerade Stau ist. Besonders viele verschiedenartige Datensätze besitzen riesige Datensammler wie Acxiom, Google, Facebook oder Geheimdienste.

III. VORHERSAGEN

Das Leben in der Datengesellschaft verändert unsere Weltsicht. Selbst die Entdeckung des Higgs-Teilchens beruhte auf der Auswertung immenser Datenmengen von physikalischen Messungen. Doch im Kern geht es bei der Massendatenauswertung um Prognosen. Diese beruhen aber nicht auf Kausalketten, wie sie der Mensch gerne formuliert, sondern auf Algorithmen, das heisst mathematischen Verfahren zur Mustererkennung in riesigen Datenmengen. Heraus kommen Wahrscheinlichkeiten – zum Beispiel, wie sich Preise entwickeln oder ob jener eilige Passant bald eine Bank überfallen wird.

- ! Nur 15% aller Daten sind strukturiert und sauber, also in Tabellen sortierbar. Der Rest besteht aus unstrukturierten Daten wie E-Mails oder Ebay-Geboten.
- ! Die meisten Datensätze können aus rechtlichen Gründen wie Datenschutz- oder Lizenzfragen nicht miteinander verknüpft werden.
- ! Um den automatischen Datenaustausch zwischen Maschinen – «Internet der Dinge» – zu ermöglichen, werden derzeit fast 340 Sextillionen neue Webadressen geschaffen.
- ! Auswertungsverfahren von grossen Datenmengen nutzt jeder von uns täglich. Sie stecken hinter Spamfiltern, Paarfindungs-Websites und der Autokorrektur im Mobiltelefon.
- ! Big-Data-Prognosemethoden beruhen auf der anhaltenden allgemeinen Erhältlichkeit persönlicher Daten.
- ! Es gibt nur rund 400 bekannte statistische Regressionsmodelle, also Verfahren, um systematische Zusammenhänge aus Datenmengen zu ermitteln.



FLASH CRASH
Blitzartiger Aktienmarkteinbruch, bei dem ein fehlerhafter Algorithmus am 6. Mai 2010 in 90 Minuten 6 % des Wertes des US-Leitindex S&P 500 vernichtete.

BRENNENDE BRONX
Ein Versuch in den 1970er-Jahren, mathematisch die optimale Lage von Feuerwehrrstationen zu bestimmen, begünstigte anhaltende Brände in der South Bronx, die 40% der Gebäude des New Yorker Stadtteils vernichteten.

KANZLERÜBERWACHUNG
Automatisierte US-Überwachungsprogramme haben Staatsführer unter anderem in Deutschland und Brasilien abgehört – und diplomatische Krisen ausgelöst.

DAS GROSSE RAUSCHEN

DIE NEUEN MASSENDATENAUSWERTUNGSMETHODEN VERSAGEN

Die angebliche Erkenntnisrevolution bleibt aus. Fehlprogrammierte Algorithmen führen zu diversen spektakulären Pannen. Auch die versprochene neue Transparenz bleibt aus. Weil das Wachstum der Verarbeitungskapazitäten von Daten nicht mit dem Wachstum der Speicherkapazitäten mithalten kann und dieses wiederum nicht ausreicht, um alle erhobenen Daten zu speichern, verschlechtert sich relativ gesehen unser Erkenntnisstand. Wir sehen hilflos zu, wie die Datenflut auf uns zukommt. Zudem führt die Analyse von immer mehr Datensätzen nicht

zwangsläufig zu stets präziseren, sondern aufgrund des Phänomens des sogenannten «Overfittings» zum Teil sogar zu schlechteren Prognosen. Auch begünstigt die Verfügbarkeit endloser unsauberer Datensätze die «statistische» Unterstützung unrichtiger Behauptungen – und führt weg von der angestrebten «Objektivierung». Aufgrund des zunehmenden Wissens um die Erhebung von Daten könnten Menschen zudem immer mehr persönliche Daten zurückhalten oder ihr Verhalten ändern, um Messresultate zu verfälschen.

IV. HANDLUNGSFELDER FÜR DAS ÜBERLEBEN IN DER DATENWELT:

DEM EIGENEN VERSTAND VERTRAUEN. ERFAHRUNGEN SAMMELN UND SELBST ENTSCHEIDEN.

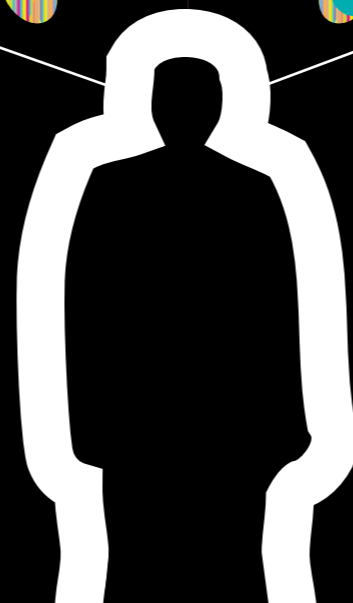
ALGORITHMEN NUR DORT ANWENDEN, WO SIE EINEN MEHRWERT SCHAFFEN. FAKTEN KRITISCH PRÜFEN.

ALGORITHMEN TRANSPARENT MACHEN.

FÄHIGKEIT ZUR REDUKTION FÖRDERN. MEHR PAUSEN MACHEN.

Auch wenn durch die wachsenden Datenmengen der tatsächliche Erkenntnisstand über die Lage der Welt sowie der sozialen und technischen Systeme zunimmt, könnte die Empfindung einer zunehmenden Intransparenz überwiegen – bedingt durch die Zunahme des Unwissens über die stets wachsenden, unverarbeiteten Datenbestände. Doch gerade dies könnte uns nutzen. Denn es zwingt uns, nach wie vor selbst zu denken, statt Entscheidungen einfach an eine letztlich eben doch nicht so intelligente Umwelt abzugeben: Im Nebel ergreift der Kapitän das Steuer selber. Zudem wäre totale Transparenz ein soziales Problem. Unsere Gesellschaft

SELBER DENKEN, KOMPLEXITÄT



REDUZIEREN, ALGORITHMEN BEWUSST NUTZEN!

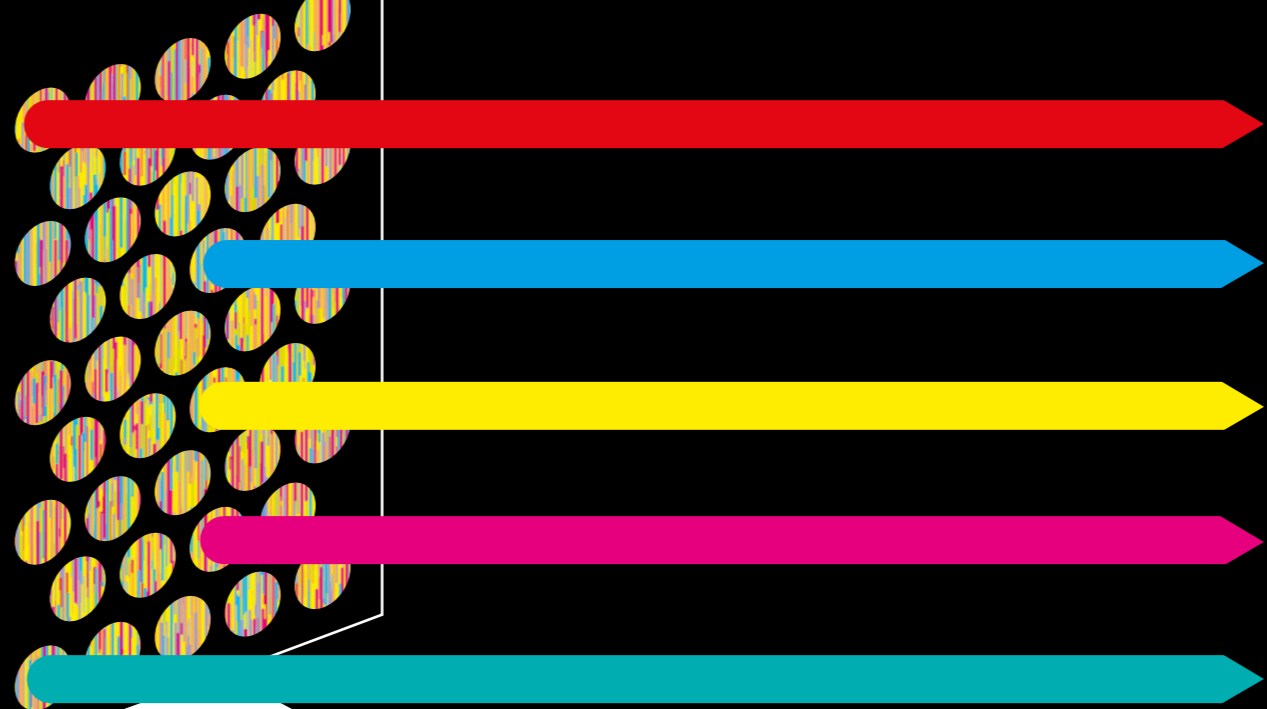
beruht auf versteckten und codierten sozialen Interaktionen. Würden diese umfassend sichtbar, müsste die Gesellschaft neu organisiert werden. Denn der Verlust der Privatsphäre, in der auch mal über den Nachbar gelästert werden darf, führte zum Verlust wertvoller sozialer Ventile. Gleichzeitig bestünde in einer algorithmisch gesteuerten Gesellschaft die Gefahr einer Normierung und des Verlustes an Vielfalt. Gerade Vielfalt hat sich in biologischen Systemen jedoch immer als Stabilisator und als Risikoabsicherung, sowie in Forschung, Wirtschaft und Politik als Grundlage für Innovation erwiesen.

DAS ZEITALTER DER TRANSPARENZ

DIE VERSPRECHUNGEN DER DATENWELT GEHEN IN ERFÜLLUNG

Eine neue Klarheit macht sich breit. Das zunehmende, datenbasierte Wissen über den Stand vieler Vorgänge in der Welt führt zu einer wachsenden Objektivität von Entscheidungsgrundlagen. Auch persönliche Leistung lohnt sich immer mehr, da diese erkannt wird. Alles wird einfacher – vom Shopping bis zur Lebensplanung. Die Zukunft ist sichtbar, weil die Folgen unserer Handlungen in Echtzeit errechnet werden können. Teilweise können die Entscheidungen über optimale Verhaltensweisen sogar komplett an Rechner abgegeben werden. Unsere Köpfe wer-

den frei für wichtigere Fragen. Zudem eröffnen sich gänzlich neue Wirtschaftsfelder, weil ultrapersonalisierte Produkte möglich werden. Wir bewegen uns durch eine mit Sensoren ausgestattete Umwelt, die intelligent auf uns reagiert, das heisst unsere Bedürfnisse erkennt und so gut wie möglich sofort erfüllt, sowie für eine kriminalitätsfreie Umgebung sorgt. Zudem werden durch die enormen Effizienzsteigerungen im Zuge der neuen, intelligenteren Verwendung von Ressourcen viele Umweltprobleme gelöst.



MEDIZINISCHE DIAGNOSEN
Nicht Ärzte, sondern eine Software hat bei der Auswertung grosser Datenmengen entdeckt, dass Babys mit regelmässigem Herzschlag später mit höherer Wahrscheinlichkeit erkranken als Babys mit unregelmässigem Herzschlag.

VORAUSSCHAUENDE VERBRECHENSBEKÄMPFUNG
Abnahme der Kriminalität durch datenbasierte Risikoschätzungen, die Einsatzkräften helfen, sich optimal zu positionieren, um Verbrechen vorzubeugen.

VORHERSAGEN
Ein Weltsimulator der ETH Zürich soll – als eine Art Internet der Daten, bei dem jeder analog zu Google Fragen zu aktuellen Zuständen der Welt sowie wahrscheinlichen künftigen Entwicklungen eingeben kann – einst helfen, Wirtschaftskrisen und Kriege vorherzusehen – und sie zu verhindern.

ROBOTERJOURNALISTEN
Das Unternehmen Narrative Science entwickelte Algorithmen, die aus Statistiken oder Twitterfeeds Informationen filtern und darauf basierend in einfachen Sätzen über Börsenkurse oder lokale Fussballspiele berichten.

VERBESSERTE KATASTROPHENHILFE
Das Fraunhofer Institut entwickelt zurzeit ein Instrument, um die Verteilung von Lieferungen bei der Katastrophenhilfe zu optimieren – auf Basis der von Opfern übermittelten Nachrichten auf Mobilgeräten und in sozialen Medien.

INNOVATION AUF REDUKTION VON KOMPLEXITÄT AUSRICHTEN: WENIGER IST MEHR.

SICH DEM MARKTWERT DER PERSÖNLICHEN DATEN BEWUSST WERDEN UND DIESE ENTSPRECHEND VERWALTEN.

DIE VORTEILE DER DATENWELT IN AUSGEWÄHLTEN BEREICHEN NUTZEN UND DORT VOLL EINTAUCHEN INS DATENMEER.

SIMPLE TECHNOLOGIE BEVORZUGEN.