

Schöne Neue Datenwelt:

Daten, Wissen, Werte & Gesellschaft

Ringvorlesung „Digitale Information und Manipulation“

27. April 2017 | HAW Hamburg

Judith Simon

Professorin für Ethik in der Informationstechnologie

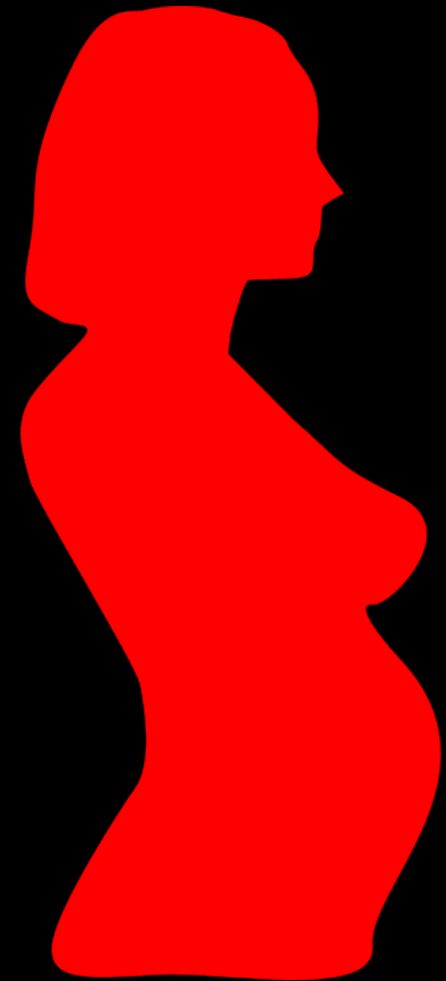
Universität Hamburg

judith.simon@informatik.uni-hamburg.de

- Big Data Basics
 - Warum & Wie?
 - Was?
 - Wer? Wirtschaft, Wissenschaft, Politik
- Big Data: Wissen, Werte & Gesellschaft
- Fazit

Big Data: Warum & Wie?

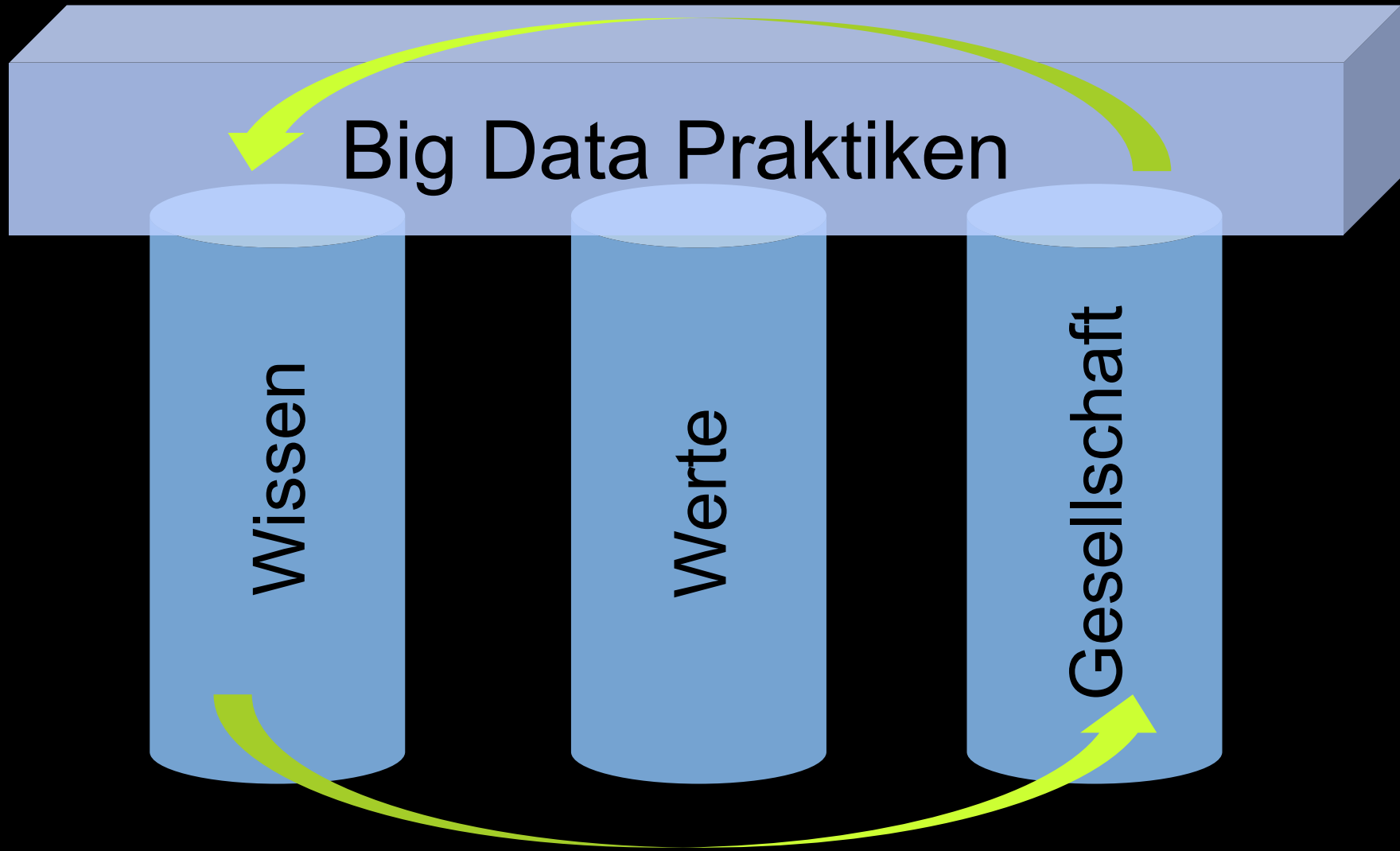
Big Data: Warum & Wie?



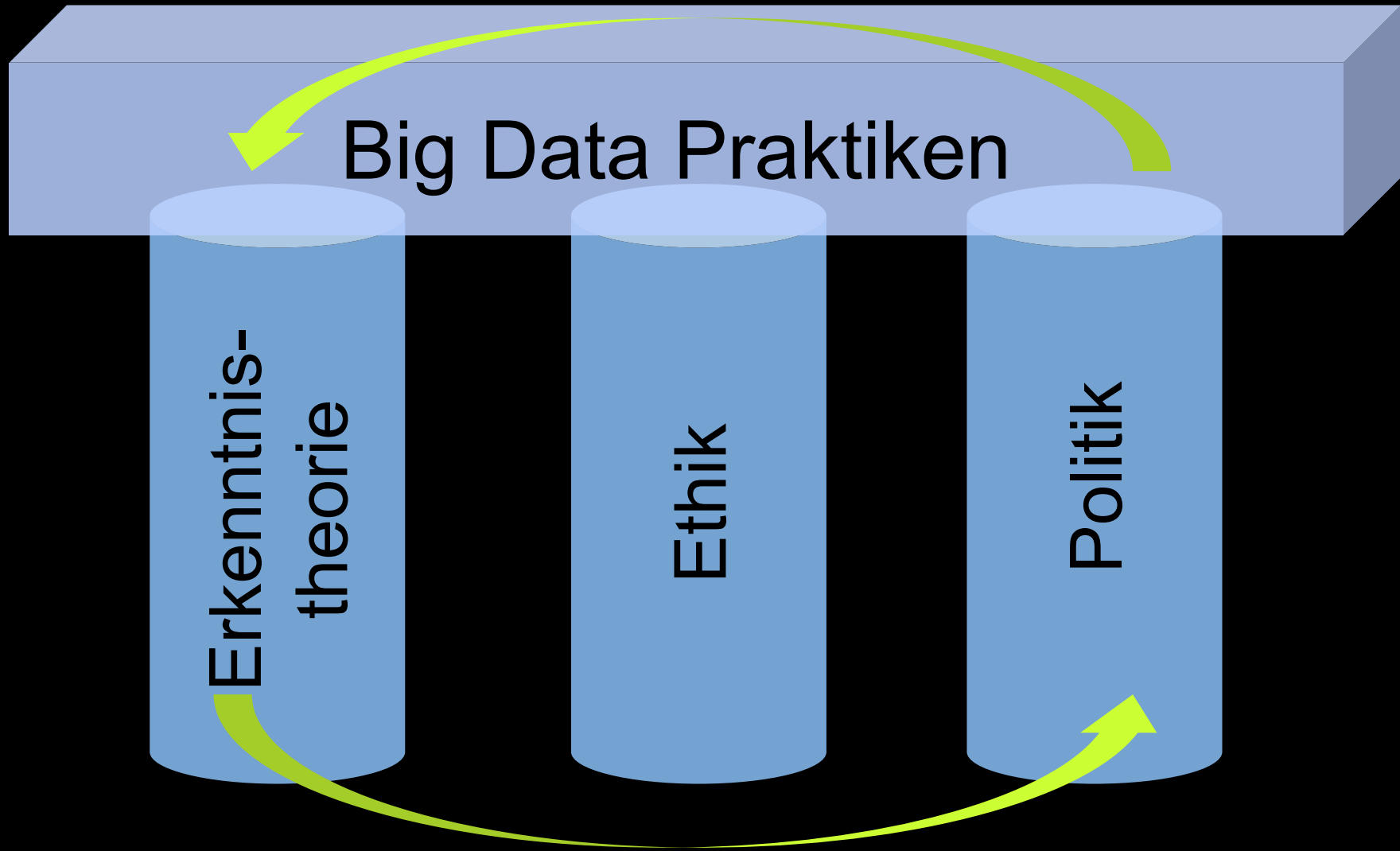
Wo liegt das Problem?

- Verletzung der Privatsphäre?
 - Illegitime Datensammlung versus informierte Zustimmung zur Nutzung der Kundeninformation?
 - Privatsphäreverletzung nicht aufgrund der Sammlung der Daten, sondern deren Verarbeitung und der Schlussfolgerungen
 - Big Data Praktiken als epistemische Praktiken, d.h. Wissenspraktiken, welche ethische und politische Implikationen haben
 - Notwendigkeit detaillierter, interdisziplinärer Analysen von Datenpraktiken in verschiedenen Sektoren

Big Data: Warum & Wie?



Big Data: Warum & Wie?



Big Data: Warum & Wie?

Big Data Praktiken

Erkenntnis-
theorie

Ethik

Politik

Recht

Ökonomie

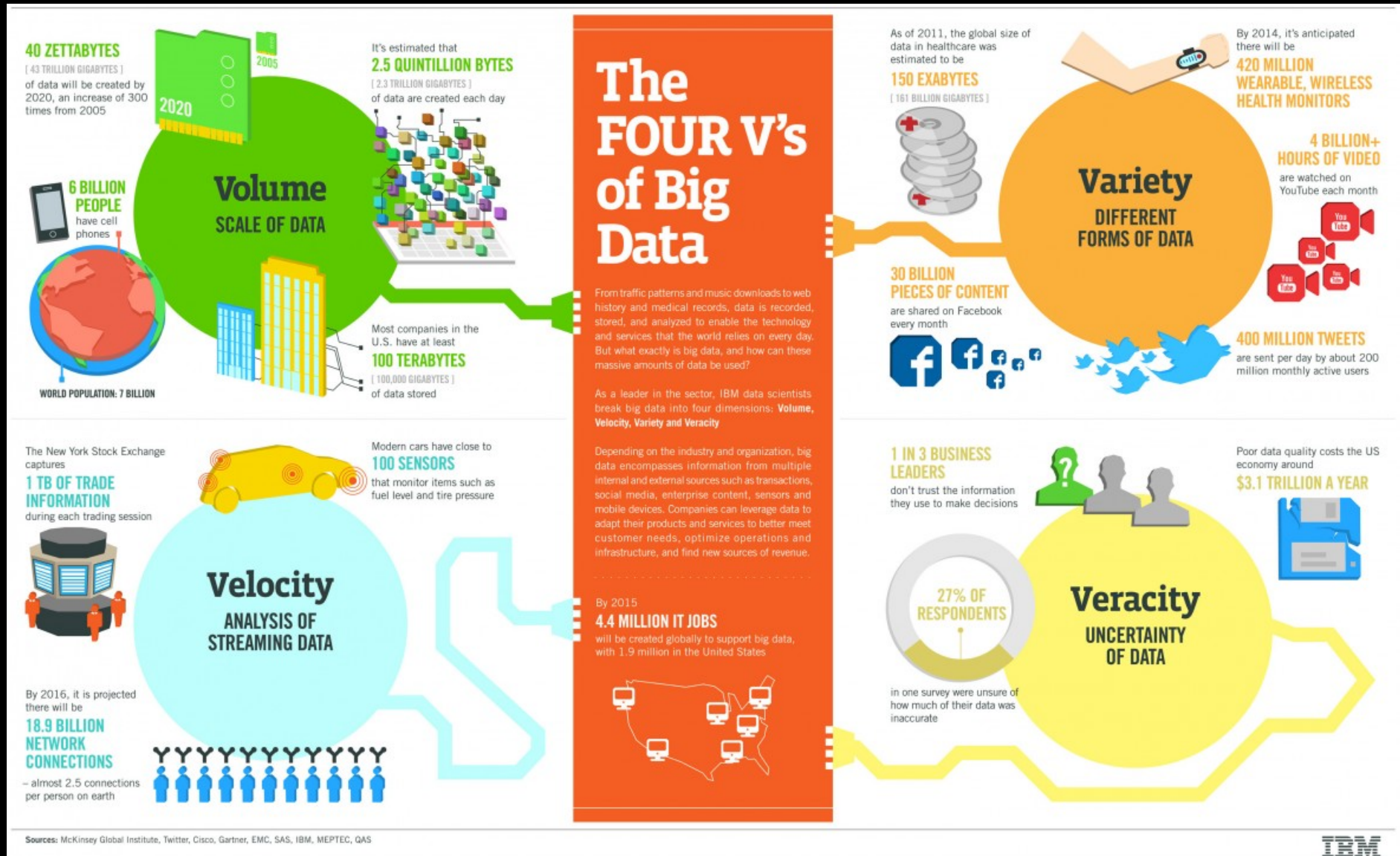
...

Big Data: Was?

Big Data: Was?

Was ist Big Data?

Big Data: Was?



Was ist Big Data?

„We define Big Data as a cultural, technological, and scholarly phenomenon that rests on the interplay of:

- 1) Technology: maximizing computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets.
- 2) Analysis: drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims.
- 3) Mythology: the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the **aura of truth, objectivity, and accuracy.**“

(Boyd & Crawford 2012:665)

(Big) Data: Arten

- Soziale Medien: Nutzer/Nutzungsdaten
 - Explizit: Kommentare, Likes, Suchanfragen, Uploads (Fotos, Videos), ...
 - Implizit: Ort, Zeit, Views, Clicks, Verweildauer,...
- Transaktionsdaten
- Ort-und zeitbezogene (Meta-) Daten
- Sensordaten, Internet der Dinge

(Big) Data: Arten

- Wissenschaftliche Daten
 - Astronomie, Physik, Biologie, Psychologie, Ökonomie, Soziologie, Pädagogik, Medizin ...
- Medizinische Daten
 - Daten aus Versuchsreihen, genetische Daten
 - Daten aus elektronischen Patientenakten, administrative Klinikdaten, ..
 - Sensordaten aus Kliniken und Pflegeeinrichtungen
 - Personal health monitoring: Wearables , Sensoren, Apps, Suchanfragen, Onlineverhalten, ...

(Big) Data: Arten

- Bürgendaten/ (Open) Government Data
 - Geburtsregister: Familienstand, Geburten, Todesfälle, ...
 - Adressen, Beruf, Sozialversicherungsnummer, ...
 - Volkszählungen
 - Finanzdaten

Big Data: Was?

A vast field of galaxies, including spiral, elliptical, and irregular shapes, scattered across a dark cosmic background. The galaxies vary in size, color (from yellow to blue), and orientation, creating a complex and dense visual representation of a large dataset.

Datenfusion

(Obsolete) Unterscheidungen

1. Sensible persönliche Daten vs. unverfängliche Daten

Big Data: Was?



(Obsolete) Unterscheidungen

1. Sensible persönliche Daten vs. unverfängliche Daten
2. Personenbezogenen Daten vs. anonyme Daten
 - Durch Aggregation und Datenverarbeitung können anonyme Daten zunehmend leicht de-anonymisiert werden
 - Diskriminierung kann auch auf Gruppenebene geschehen, Identifikation nicht zwingend notwendigerweise

Big Data: Wer?

Big Data: Wer?

„Specifically, people and organizations are divided into three categories: those who **create** data (both consciously and by leaving digital footprints), those who have the means to **collect** it, and those who have expertise to **analyze** it.“

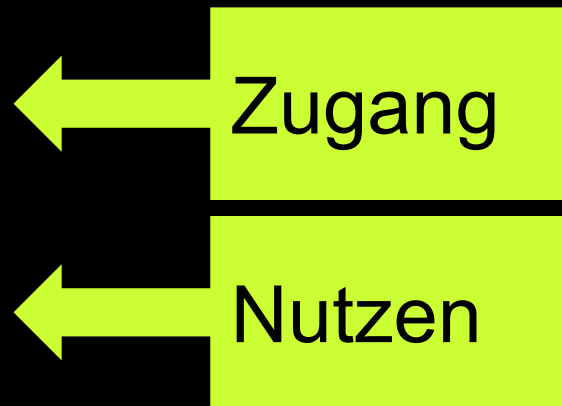
(Manovich 2011)

Big Data: Wer?

- Daten-Bereitsteller

- Daten-Sammler/Besitzer

- Daten-Analysten



Big Data: Wer?

„But these capabilities, most of which are not visible or available to the average consumer, also create an **asymmetry of power** between those who hold the data and those who intentionally or inadvertently supply it.“

White House Report on Big Data (May 2014):3

Wer sammelt (personenbezogene) Daten?

- Wissenschaft
- Staaten
- Unternehmen
- (Nutzer selbst)



facebook

Google

Wer sammelt (personenbezogene) Daten?

- Wissenschaft
- Staaten
- Unternehmen
- (Nutzer selbst)

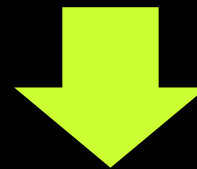


Unterschiedliche Funktionen
& Implikationen von Big Data

Wer sammelt (personenbezogene) Daten?

- Wissenschaft
- Staaten
- Unternehmen
- (Nutzer selbst)

Unterschiedliche Funktionen
& Implikationen von Big Data



Notwendigkeit differenzierter,
domainspezifischer Analysen

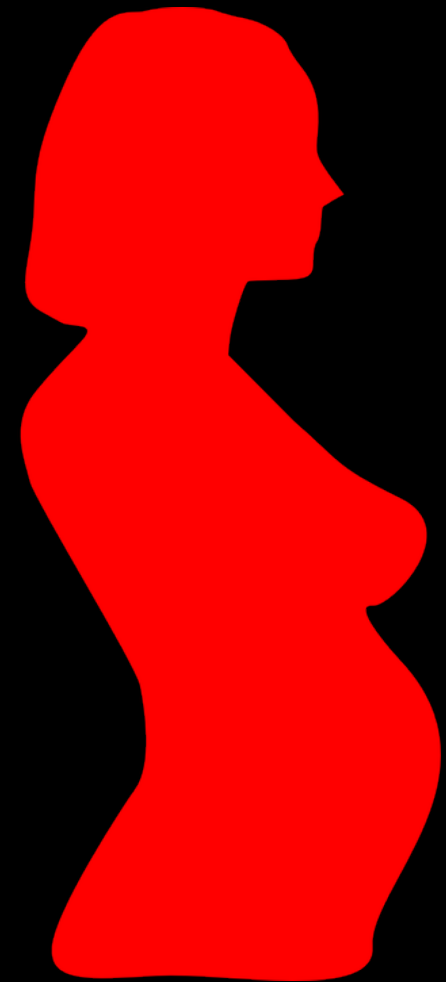
Big Data: Unternehmen



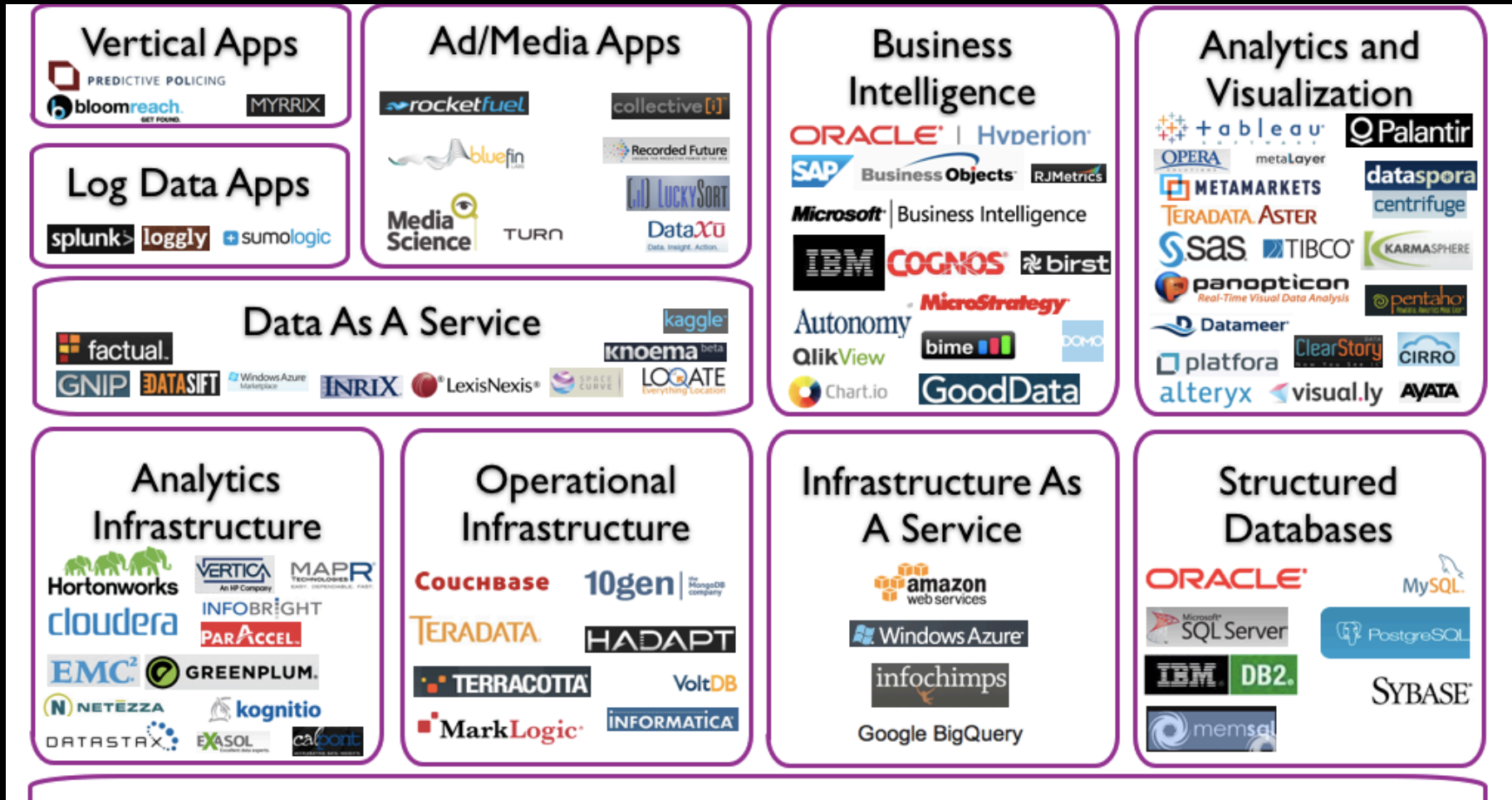
Neue Datensammler

The Google logo is displayed in its characteristic multi-colored font (blue, red, yellow, green, red) on a white background.The Facebook logo is displayed in white lowercase letters on a blue background.

Big Data: Unternehmen



Big Data: Unternehmen



Big Data: Unternehmen

- Es sind doch nur Metadaten
- Ortsdaten+ Zeitdaten+ Landkarte → enorm aufschlussreich
- Personenbasierte „Rohdaten“ + allgemeines Hintergrundwissen = angereicherte personen-basierte Daten
- + Semantische Karten + soziale Informationen + prädiktive Modelle

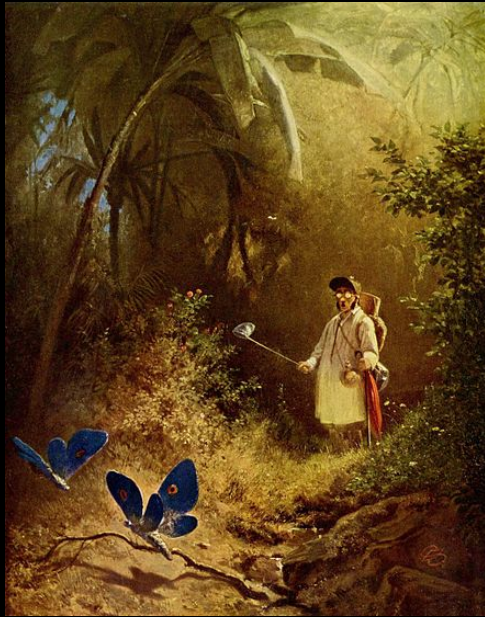


→ Privatsphäre, Diskriminierung, Autonomie, Gleichheit, etc.

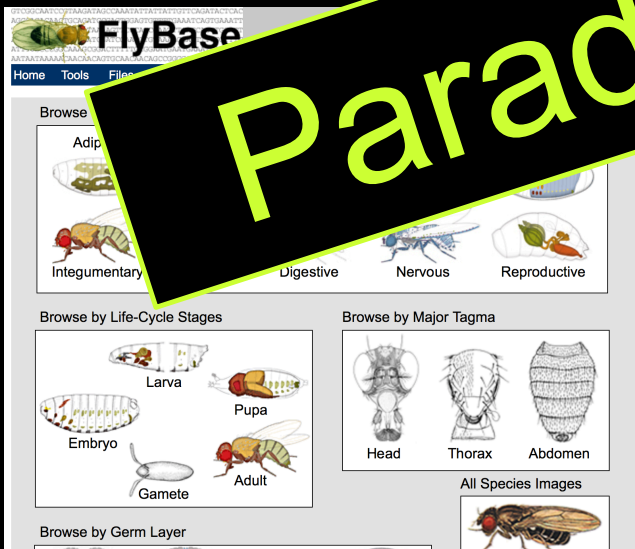
(Hoffmann 2014: „Zur Hermeneutik digitaler Daten“, Presentation at „Die Vermessung des Menschen, Konstanz, May 8 2014)

Big Data: Wissenschaft

Big Data: Wissenschaft



Paradigmenwechsel?



Big Data: Wissenschaft

CHRIS ANDERSON SCIENCE 06.23.08 12:00 PM

THE END OF THEORY: THE DATA DELUGE MAKES THE SCIENTIFIC METHOD OBSOLETE



Illustration: Marian Bantjes

"All models are wrong, but some are useful."

So proclaimed statistician George Box 30 years ago, and he was right. But what choice did we have? Only models, from cosmological equations to theories of human behavior, seemed to be able to consistently, if imperfectly, explain the world around us. Until now. Today companies like Google, which have grown up in an

Big Data: Wissenschaft

- Neuer Empirismus
 - “End of theory”: Daten können für sich selbst sprechen
 - Rein induktives Vorgehen, keine Notwendigkeit für Hypothesen oder wissenschaftliche Theorien
 - Korrelation ist wichtiger als Kausalität
 - “Aura of objectivity, truth, and accuracy” (Boyd & Crawford 2012)

CHRIS ANDERSON SCIENCE 06.23.08 12:00 PM

THE END OF THEORY: THE DATA DELUGE MAKES THE SCIENTIFIC METHOD OBSOLETE




Illustration: Marian Bantjes

“All models are wrong, but some are useful.”

So proclaimed statistician George Box 30 years ago, and he was right. But what choice did we have? Only models, from cosmological equations to theories of human behavior, seemed to be able to consistently, if imperfectly, explain the world around us. Until now. Today companies like Google, which have grown up in an

Big Data: Wissenschaft

- Neuer Empirismus
 - “End of theory”: Daten können für sich selbst sprechen
 - Rein induktives Vorgehen, keine Notwendigkeit für Hypothesen oder wissenschaftliche Theorien
 - Korrelation ist wichtiger als Kausalität
 - “Aura of objectivity, truth, and accuracy” (Boyd & Crawford 2012)
- Versus Forschungsrealitäten
 - Kitchin (2014): Big Data, new epistemologies and paradigm shifts
 - Leonelli (2014): What difference does quantity make? On the epistemology of Big Data in biology

Big Data: Wissenschaft

- Neuer Empirismus
 - “End of theory”: Daten können für sich selbst sprechen
 - Rein induktives Vorgehen, keine Notwendigkeit für Hypothesen oder wissenschaftliche Theorien
 - Korrelation ist wichtiger als Kausalität
 - “Aura of objectivity, truth, and accuracy” (Boyd & Crawford 2012)
- Versus Forschungsrealitäten
 - Daten hängen ab von: Plattformen, Ontologien, Ein- & Ausschlußkriterien, Formatierungen, wissenschaftlichen und technischen Praktiken, regulativen Rahmenbedingungen, ...
 - Datenpräparation als arbeitsintensiver Prozess mit geringem Reputationsgewinn
 - Probleme bzgl. Datenzugang und Datenkompetenz

(Leonelli 2014, Kitchin 2014)

Big Data: Wissenschaft

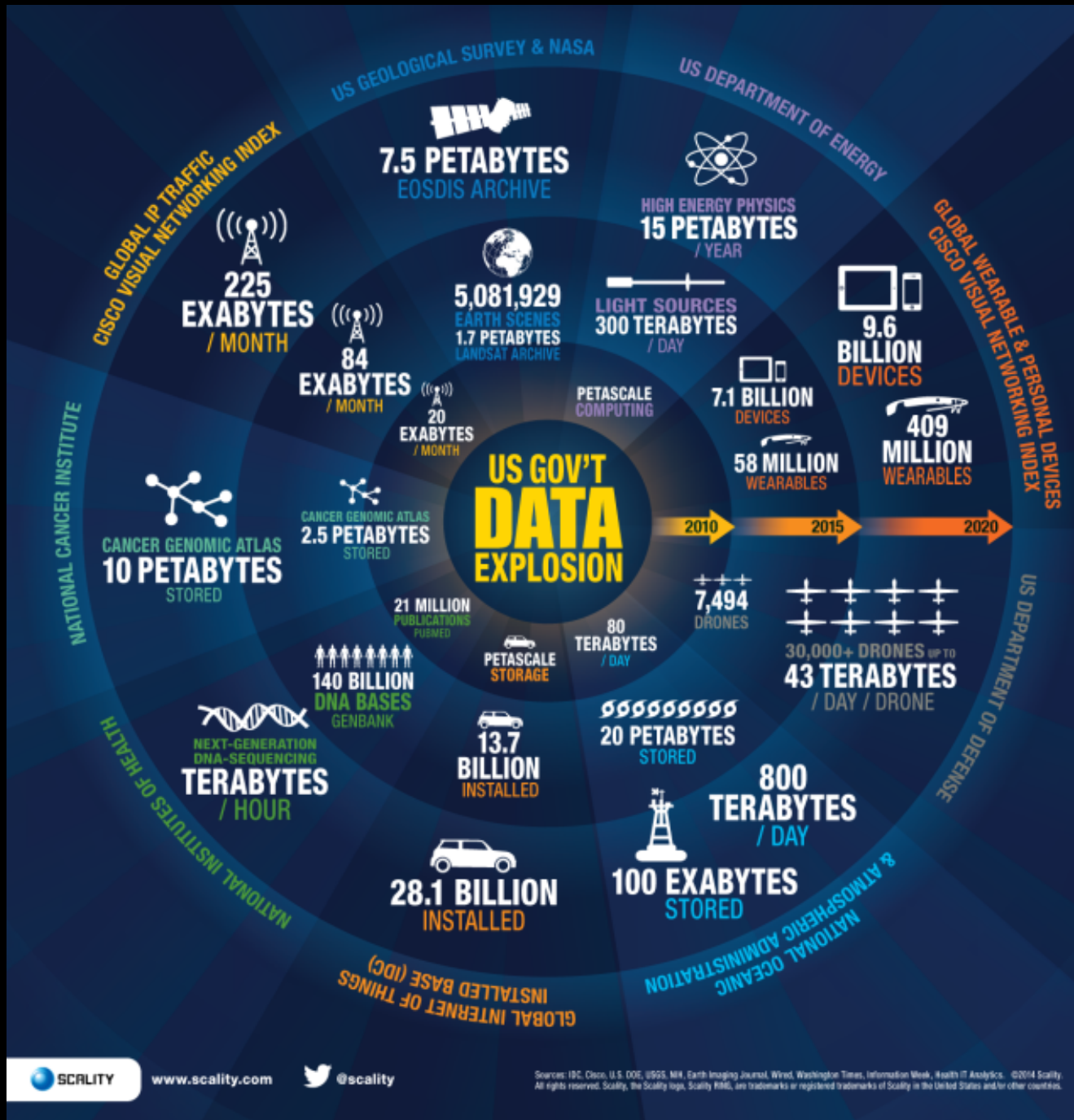
- Neuer Empirismus
 - “End of theory”: Daten können für sich selbst sprechen
 - Rein induktives Vorgehen, keine Notwendigkeit für Hypothesen oder wissenschaftliche Theorien
 - Korrelation ist wichtiger als Kausalität
 - “Aura of objectivity, truth, and accuracy” (Boyd & Crawford 2012)
- Erkenntnistheoretische Alternativen aus den Digital Humanities & Computational Social Science
 - Hybride Ansätze, die Induktion, Deduktion und Abduktion verbinden
 - Fachliche situierte und kontextualisierte Forschung

(Kitchin 2014)

Big Data: Wissenschaft

- Neuer Empirismus
 - “End of theory”: Daten können für sich selbst sprechen
 - Rein induktives Vorgehen, keine Notwendigkeit für Hypothesen oder wissenschaftliche Theorien
 - Korrelation ist wichtiger als Kausalität
 - “Aura of objectivity, truth, and accuracy” (Boyd & Crawford 2012)
- Wissenschaftlich höchst umstritten
- Rethorik verführerisch für Politik?

Big Data: Politik



www.scality.com



Sources: IDC, Cisco, U.S. DOE, ICGS, NIH, Earth Imaging Journal, Wired, Washington Times, Information Week, Health IT Analytics. ©2014 Scality. All rights reserved. Scality, the Scality logo, Scality PaaS, are trademarks or registered trademarks of Scality in the United States and/or other countries.

© <https://whatsthebigdata.com/2014/09/15/u-s-governments-data-explosion-infographic/>

Big Data: Politik



Big Data: Politik

Der Terminus „Statistik“ bezeichnet seit dem 18. Jahrhundert die systematische Sammlung demographischer und ökonomischer Daten durch den Staat.

„As the etymology of the word shows, statistics is connected with the construction of the state, with its unification and administration.“

(Desrosières 1998:8)

„The need to know a nation in order to govern it led to the organization of official bureaus of statistics (...).“ (Desrosières 1998:16)

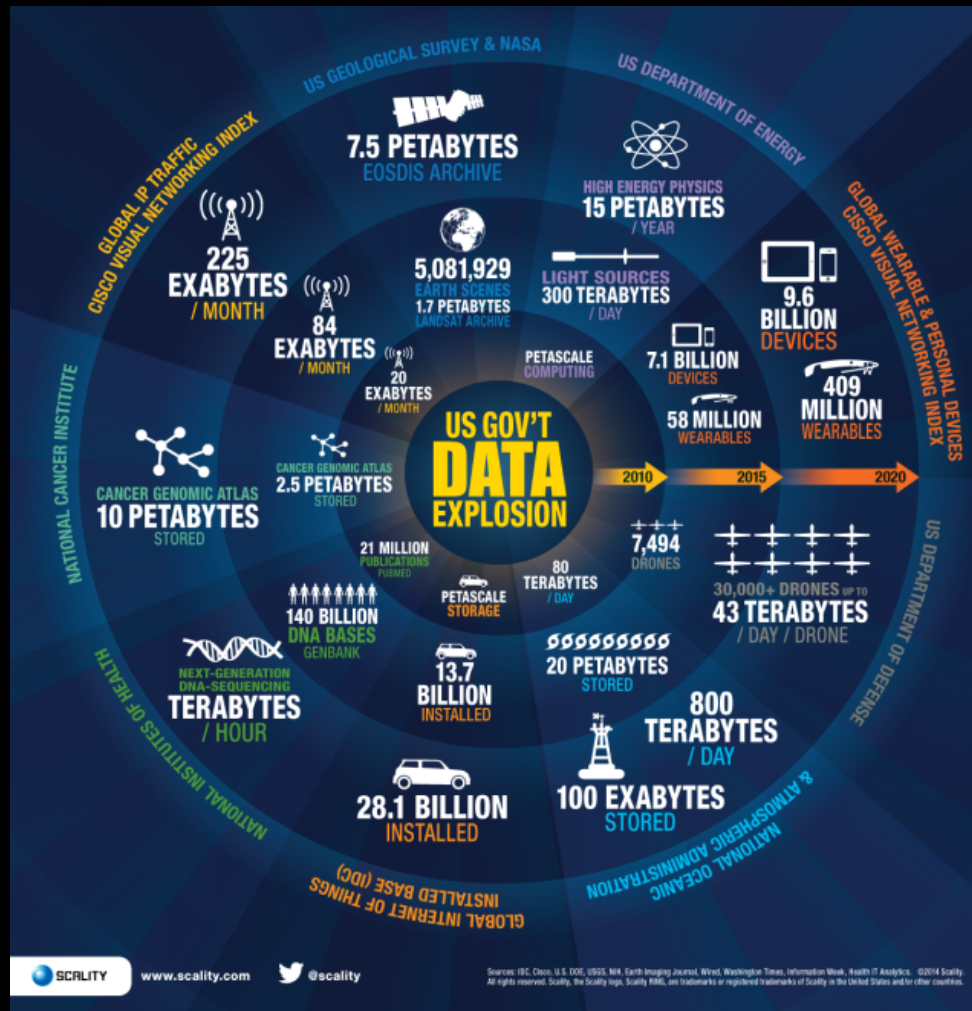
→ Geschichte von Nationalstaaten als Geschichte der Statistik, als eine Geschichte des Sammelns und Verarbeitens von Daten

→ Wissen <> Macht

Big Data: Wissen, Werte & Gesellschaft

Big Data & Politik: Wissen, Werte & Gesellschaft

- Big Data in Politik & Administration: Evidence-based policy making?



Big Data & Politik: Wissen, Werte & Gesellschaft

- Big Data in Politik & Administration: Evidence-based policy making?
- Verlangt nach epistemischer, politischer und ethischer Analyse und Begründung
 - Epistemisch: Qualität der Daten, der Ergebnisse und der Prognosen
 - Politisch: das beste Argument gewinnt nicht notwendigerweise in der Politik; umstrittene Rolle von Expertenwissen, Gefahr der Technokratie, etc.
 - Ethisch: Privatsphäre, Diskriminierung, Gleichheit, Autonomie, ..
- Probleme
 - Epistemische Rechtfertigung: schwierig aufgrund des Mangels an Zugang zu Daten und der Kompetenz in kritischer Datenanalyse
 - Darstellung von Big Data als neutraler, Theorie- und Verzerrungsfreier Wissenschaft erschwert kritische Analyse ggf. zusätzlich

Big Data & Politik: Wissen, Werte & Gesellschaft



MOTHERBOARD

TRUMPING

The Data That Turned the World Upside Down

HG Hannes Grassegger and Mikael Krogerus
Jan 28 2017, 3:15pm

Psychologist Michal Kosinski developed a method to analyze people in minute detail based on their Facebook activity. Did a similar tool help propel Donald Trump to victory? Two reporters from Zurich-based Das Magazin went data-gathering.

Big Data & Politik: Wissen, Werte & Gesellschaft


LIBERAL ⓘ

SHOWING POSTS ABOUT:
"PRESIDENT OBAMA"

CONSERVATIVE ⓘ

Reverb Press ✓
8 hours ago


Obama is a tough act to follow.



Watch Chuck Todd OBLITERATE Kellyanne Co...
Meet The Press host Chuck Todd was ready for Kellyanne Conway whe...
REVERBPRESS.COM | BY MEGAN HAMILTON

Stand Up, Speak Out, Fight Back ✓
7 hours ago

What the hell!



Obama Quietly SHUT Down Our Aerial Surveill...
Obama is on his way out the doors of the White House, and it just feels ...
IHAVETHETRUTH.COM

<http://graphics.wsj.com/blue-feed-red-feed/#/obama>

Big Data & Politik: Wissen, Werte & Gesellschaft

Social Media

Die dunkle Macht der Algorithmen

von Stefan Betschon / 24.11.2016, 05:30 Uhr

In den Social Media wird die politische Kommunikation durch automatisch generierte Fake News geprägt. Jede fünfte Nachricht in den Social Media wurde algorithmisch, durch Bots, produziert.

Bots and Automation over Twitter during the U.S. Election

[COMPROP](#) DATA MEMO 2016.4 / 17 NOV 2016

Bence Kollanyi
Corvinus University
kollanyi@gmail.com
[@bencekollanyi](https://twitter.com/bencekollanyi)

Philip N. Howard
Oxford University
philip.howard@oii.ox.ac.uk
[@pnhoward](https://twitter.com/pnhoward)

Samuel C. Woolley
University of Washington
samwooll@uw.edu
[@samuelwoolley](https://twitter.com/samuelwoolley)

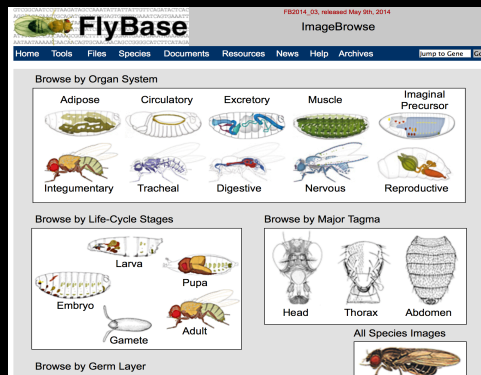
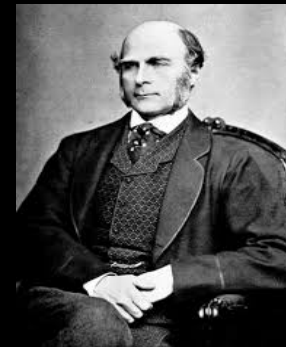
ABSTRACT

Bots are social media accounts that automate interaction with other users, and political bots have been particularly active on public policy issues, political crises, and elections. We collected data on bot activity using the major hashtags related to the U.S. Presidential Election. We find that that political bot activity reached an all-time high for the 2016 campaign. (1) Not only did the pace of highly automated pro-Trump activity increase over time, but the gap between highly automated pro-Trump and pro-Clinton activity widened from 4:1 during the first debate to 5:1 by election day. (2) The use of automated accounts was deliberate and strategic throughout the election, most clearly with pro-Trump campaigners and programmers who carefully adjusted the timing of content production during the debates, strategically colonized pro-Clinton hashtags, and then disabled activities after Election Day.

Big Data & Wissenschaft: Wissen, Werte & Gesellschaft

Big Data & Wissenschaft: Wissen, Werte & Gesellschaft

- Big Data in der Wissenschaft



Big Data & Wissenschaft: Wissen, Werte & Gesellschaft

- **Wissen:**
 - Qualität der Daten, Verfahren und Schlussfolgerungen, ...
- **Werte:**
 - Veränderung wissenschaftlicher Werte durch Big Data (Einfachheit, Robustheit, ...)
 - Veränderung gesellschaftlicher Werte durch Wissenschaft
 - Gesellschaftlicher Wert von Wissenschaft
- **Gesellschaft**
 - Politischer Ge/Missbrauch wissenschaftlicher Erkenntnisse oder deren Ignorieren

Big Data & Wirtschaft: Wissen, Werte & Gesellschaft

Big Data & Wirtschaft: Wissen, Werte & Gesellschaft

- Wissen:
 - Qualität der Daten, Verfahren und Schlussfolgerungen, ...
- Werte:
 - Anwendung von Big Data in verschiedenen Sektoren und Auswirkungen auf zentrale Werte wie Gleichheit, Autonomie, Gerechtigkeit, Freiheit, Privatsphäre, ...
- Gesellschaft:
 - Divergente Interessen an Big Data: von nachrichtentechnischer und wirtschaftlicher Nutzung zu datenschutzrechtlicher oder kartellrechtlicher Regulierung ...
 - Möglichkeiten politischer, rechtlicher & gesellschaftlicher Intervention
 - ...

Fazit

Big Data: Was Nun?

- Big Data für Politik & Forschung → Forschung zu & Politik/Governance für Big Data
 - Wie erhöht man die Vertrauenswürdigkeit von Datenpraktiken?
 - Wie kann man epistemische, politische und ethische Überprüfbarkeit unterstützen?

Big Data Governance

Recht

Selbst-
regulierung

Governance
By Design

Bildung

Big Data: Was Nun?

- Big Data für Politik & Forschung → Forschung zu & Politik/Governance für Big Data
 - Wie erhöht man die Vertrauenswürdigkeit von Datenpraktiken?
 - Wie kann man epistemische, politische und ethische Überprüfbarkeit unterstützen?

Big Data Governance

Recht

Selbst-
regulierung

Governance
By Design

Bildung

Drei Argumentationslinien zur Regulierung

1. Verteidige den Status Quo

- Es bedarf keiner weiteren Regulation, Markt reguliert sich selbst

2. Verbesserte Regulation

- Suche nach und Implementierung effektiverer Mittel um große Internetfirmen zu regulieren

3. Veränderung der ökonomischen Grundlagen von Datenmärkten

- Alternative Infrastrukturen
- Datenbesitz: Nutzer als Besitzer, Daten als öffentliches Gut, ...

Big Data: Was Nun?

- Big Data für Politik & Forschung → Forschung zu & Politik/Governance für Big Data
 - Wie erhöht man die Vertrauenswürdigkeit von Datenpraktiken?
 - Wie kann man epistemische, politische und ethische Überprüfbarkeit unterstützen?

Big Data Governance

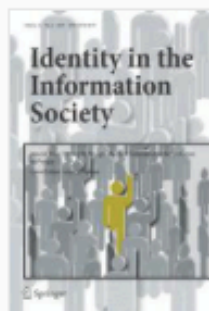
Recht

Selbst-
regulierung

Governance
By Design

Bildung

Big Data: Was Nun?



[Identity in the Information Society](#)

August 2010, Volume 3, [Issue 2](#), pp 267–274

Privacy by Design

Authors

[Authors and affiliations](#)

Peter Schaar 

[Open Access](#) | Article

First Online: 01 April 2010

DOI: 10.1007/s12394-010-0055-x

Cite this article as:

Schaar, P. IDIS (2010) 3: 267.

doi:10.1007/s12394-010-0055-x

18

Citations

4.6k

Downloads

Abstract

In view of rapid and dramatic technological change, it is important to take the special requirements of privacy protection into account early on, because new technological systems often contain hidden dangers which are very difficult to overcome after the basic design has been worked out. So it makes all the more sense to identify and examine possible data

Big Data: Was Nun?

Vrije Universiteit Brussel

From the Selected Works of Mireille Hildebrandt

2011

Legal Protection by Design. Objections and Refutations

Mireille Hildebrandt, *Radboud University Nijmegen*

Big Data: Was Nun?

- Values in Design/Value-Sensitive Design



<http://www.nyu.edu/projects/nissenbaum/>



<https://ischool.uw.edu/people/faculty/batya>

Big Data: Was Nun?

- Big Data für Politik & Forschung → Forschung zu & Politik/Governance für Big Data
 - Wie erhöht man die Vertrauenswürdigkeit von Datenpraktiken?
 - Wie kann man epistemische, politische und ethische Überprüfbarkeit unterstützen?

Big Data Governance

Recht

Selbst-
regulierung

Governance
By Design

Bildung

Big Data: Was Nun?

Customers Who Bought This Item Also Bought



Die digitale Bildungsrevolution: Der radikale Wandel des...
> Jörg Dräger
★★★★☆ 17
Gebundene Ausgabe
EUR 17,99 ✓Prime



Big Data: Die Revolution, die unser Leben verändern wird
> Viktor...
★★★★☆ 14
Gebundene Ausgabe
EUR 24,99 ✓Prime



Die Lüge der digitalen Bildung: Warum unsere Kinder das Lernen...
> Gerald Lembke
★★★★☆ 6
Gebundene Ausgabe
EUR 19,99 ✓Prime



Big Data: A Revolution That Will Transform How We Live, Work and Think
> Viktor...
★★★★☆ 18
Taschenbuch
EUR 7,99 ✓Prime



**Vielen Dank für Ihre
Aufmerksamkeit!**

Judith Simon
Universität Hamburg
simon@informatik.uni-hamburg.de