

Big Data Praktiken & die Verschränkung von Epistemologie, Ethik und Politik

Entscheidende Daten: Die Vermessung der Gesellschaft

Konstanz | 8-9. Mai 2014

Judith Simon

Institute of Philosophy | University of Vienna

Technologies in Practice Group | IT University Copenhagen

Der Supermarkt und das Mädchen



dreamsbüro.com

Der Supermarkt und das Mädchen

Wo liegt das Problem?

Wo liegt das Problem?

- Verletzung der Privatsphäre?
 - Illegitime Datensammlung versus informed consent 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

Was ist Big Data?

Was ist Big Data?

- Volume
- Velocity
- Variety
- (Complexity)

Was ist Big Data?

- Mehr
- Unscharf
- Korrelation

Wer sammelt (personenbezogene) Daten?

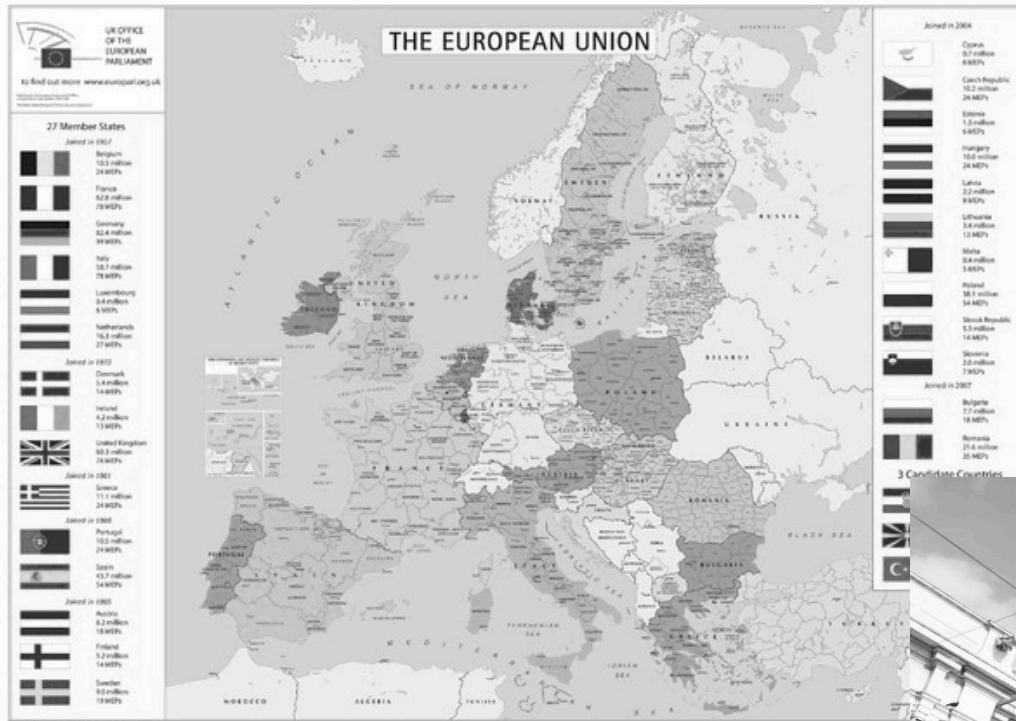
- Staaten
- Unternehmen
- Wissenschaft
- Nutzer selbst



facebook

Google

Datensammler



Datensammler

- „By the 18th century, the term "statistics" designated the systematic collection of demographic and economic data by states.“
(http://en.wikipedia.org/wiki/History_of_statistics)
 - „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 der Staatenentwicklung als Geschichte des Datensammelns, der Statistik
- Wissen <> Macht

Datensammler



From: Tafeln zur Statistik der oesterreichischen Monarchie (1839)

Religions-Verschiedenheit.

Länder	Katholiken	Unite	Nicht-unite	Protestanten			Andere Sekten	Juden	Bevölkerung (ohne Militär)
				Anglikaner	Evangelischer	Unitarier			
Oesterreich unter der Enns	1,319,824	700	250	11,000	1,800		30	2,800	1,336,404
Oesterreich ob der Enns	874,902			15,900	20		40		842,262
Steiermark	940,859			5,900	80				946,239
Kärnten und Krain	727,256	200	200	17,800			10		745,466
Küstenland	467,009		1,700	350	450		4,150		472,659
Triest	825,310				120		960		826,390
Böhmen	3,918,982			28,500	51,800		70,700		4,069,982
Mähren und Schlesien	1,967,729			79,400	29,200		37,200		2,113,729
Galizien	3,032,863	3,041,700	246,700	23,600	2,500		1,700	265,800	4,634,865
Dalmatien	310,176	658	72,634		26			560	384,104
<i>Summe</i>	13,754,900	3,042,258	321,534	181,250	86,096		1,770	382,280	16,371,038
Lombardie	2,494,410			157	401			2,897	2,497,895
Venedig	2,114,178	57	385	157	45			4,384	2,119,206
<i>Summe</i>	4,608,588	57	385	314	446			7,281	4,617,101
<i>Hauptsumme</i>	17,363,488	3,042,315	321,919	181,564	86,542			389,561	20,988,139
Ungarn	6,327,400	910,100	1,288,900	849,600	1,781,800			259,200	11,917,100
Siebenbürgen	208,000	562,900	678,200	214,400	339,400	42,300		3,500	2,043,700
Militärgränzen	470,379	56,909	363,101	12,656	31,176	2,610		464	1,137,289
<i>Summe</i>	7,505,779	1,529,909	2,530,201	1,076,656	2,152,376	44,910		263,264	15,103,089
<i>Totale</i>	25,469,267	3,572,218	3,852,120	1,258,250	2,298,918	44,910	1,770	652,825	36,091,275

Datensammler

„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 (Mai 2014):3

Neue Datensammler



	Requires a warrant for content	Tells users about government data requests	Publishes transparency reports	Publishes law enforcement guidelines	Fights for users' privacy rights in courts	Fights for users' privacy rights in Congress
amazon	★	★	★	★	★	★
Apple	★	★	★	★	★	★
at&t	★	★	★	★	★	★
comcast	★	★	★	★	★	★
Dropbox	★	★	★	★	★	★
facebook	★	★	★	★	★	★
foursquare	★	★	★	★	★	★
Google	★	★	★	★	★	★

Datensammler



Google

facebook

Big Data Praktiken als Wissenspraktiken

Was wird gesammelt?

Big Data Praktiken als Wissenspraktiken

- Datenformen:
 - Bürgerdaten
 - Familienregister: Familienstand & Beziehungen, Geburtsort & -tag, ...
 - Adresse, Beruf, Einkommen, Finanz- und Steuerdaten, Sozialversicherungsnummer,...
 - Persönliche Informationen (Tavani 2005):
 - nonpublic personal information: medizinische, finanzielle Informationen, ...
 - public personal information: Arbeitsplatz, Schule, Auto, Adresse, ...
 - Nutzungsinformation:
 - Explizit: Kommentare, Likes, Suchanfragen, Uploads (Photos, Videos), ...
 - Implizit: Ort, Views, Clicks, Zeit online, Transaktionsdaten,
 - Sensordaten & Internet der Dinge

(Hinfällige) Unterscheidungen

- Personenbezogene Information versus nicht personenbezogene Information
- Identitätsbezogene versus statistische Information
- Durch Aggregation, Verbindung & Verarbeitung können nicht-personenbezogene Daten personenbezogen werden
- Möglichkeit der Re-Identifizierung
 - Small Data Bsp: FWF Fragebogen

Nutzung

- Ökonomisch vs. Politisch
- Funktionen: Personalisierung, Identifizierung, Vorhersage, Kontrolle, ...
- Nutzen & Gefahren



Google

facebook

Überwachung, Kontrolle – und Ablenkung?



Minority Report

Überwachung, Kontrolle – und Ablenkung?



Minority Report-like adverts 'may hit the UK next year'

(Dave Lee, Feb 2011)

<http://www.bbc.co.uk/news/technology-12582477>

Big Data Praktiken: Epistemische, Ethische & Politische Desiderate

- Privatsphäre & Datenschutz
- Aufmerksamkeitsökonomie versus Aufmerksamkeitsökologie
- Zugang
 - Wer hat Zugriff auf welche Daten? Open vs. Closed
 - Was ist gerechter Zugang?
 - Konzentration von Daten bei großen Firmen & Problem der Monopolisierung
- Kompetenzen
 - Wer kann Daten nutzen? Verteilung von Kompetenzen und deren Aufbau
- Korrelation statt/versus Kausalität
 - Faktisch versus epistemisch
 - Bsp: Preisdiskriminierung (real) oder präventive Verhaftungen (fiktiv) hebeln Differenz zwischen Korrelation & Kausalität aus
-

Konklusionen

Wie kann man diesen
Herausforderungen begegnen?

Big Data & Society: Critical Interdisciplinary Inquiries

Twitter: @BigDataSoc

Blog: bigdatasoc.blogspot.com/



- **Big Data: Methoden**
 - Methodologische Innovationen, Experimente mit Datenrepräsentation, Visualisierung, Sonifizierung, Simulation, ...
- **Big Data: Epistemologien**
 - Verständnis neuer Wissensprozesse, -praktiken und –produkte im Zeitalter von Big Data; Status von Kausalität, ...
- **Big Data: Ontologien**
 - Performativität von Big Data Praktiken; Erstellen von Subjektivität, Identität, Kollektiven und Handlungsfähigkeiten, von Orten (materiell, virtuell, hybrid) und deren Relationen, ...

- Big Data: Politiken
 - Ethische und politische Implikationen von Surveillance, Tracking, Verknüpfung und Inferenzen; Crowdsourcing, Citizen Science & Neuordnung von Autoritäten, Open Access & Verteilung von Zugang und Kompetenzen, ...
- Big Data: Ökonomien
 - Wert/Wertung von Daten, Akademische Ökonomien von Big Data (Scientometrics, Indices, etc), Copyright, Open Source, ...
- Big Data: Ökologie
 - Räumliche (Verteilung) und zeitliche Aspekte (Lifecycle, Zirkulation) von Big Data

Wie kann man diesen Herausforderungen begegnen?

Big Data Governance

Hard Law

Soft Law

**Governance
By Design**

Education

Onlife-Initiative

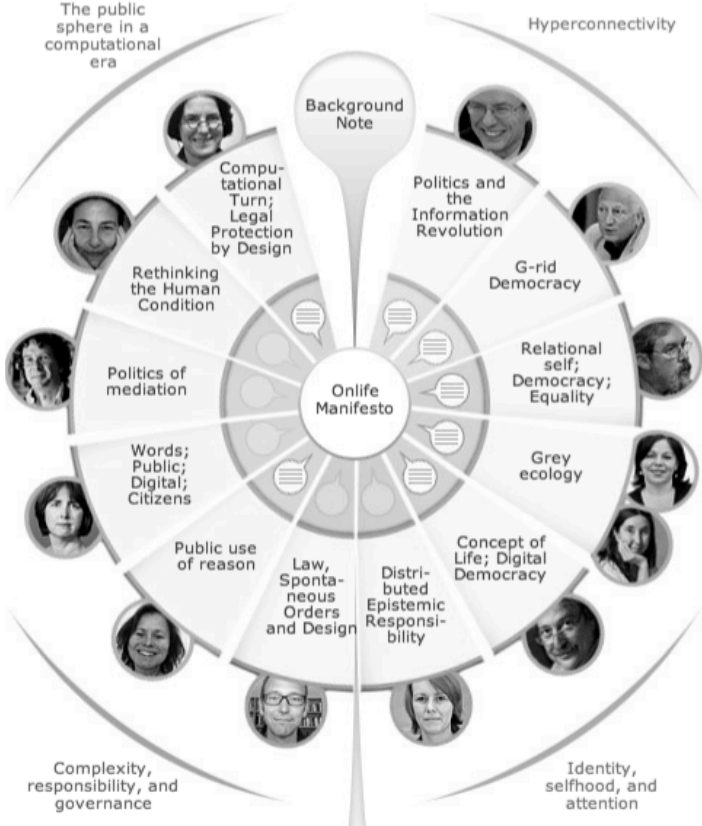
<http://ec.europa.eu/digital-agenda/en/onlife-web-output>

Digital Agenda for Europe – European Commission

ec.europa.eu/digital-agenda/en/onlife-web-output

Apple Yahoo! Google Maps YouTube Wikipedia News ▾ Beliebt ▾

- ▶ About
- ▶ Consultations
- ▶ Digital Agenda Assembly 2012
- ▶ Collective Awareness
- ▶ Digital Champions
- ▶ Digital Agenda Assembly 2011
- ▶ Going local 2011
- ▶ Big Ideas community
- ▶ Stakeholder Day 2010
- ▶ Digital Futures
 - ▶ About
 - ▶ Interviews
 - ▶ Onlife Initiative
 - ▶ Onlife Web Output
 - ▶ About Onlife
 - ▶ On Site Meetings
 - ▶ Workshop members
 - ▶ Join the Debate
 - ▶ OnLife News
 - ▶ Policy Making 3.0



The diagram is a circular infographic titled "Onlife Manifesto" at its center. It is divided into 12 segments, each containing a topic and a small circular portrait of a person. The segments are: "Background Note" (top), "Politics and the Information Revolution", "G-rid Democracy", "Relational self; Democracy; Equality", "Grey ecology", "Concept of Life; Digital Democracy", "Distributed Epistemic Responsibility", "Law, Spontaneous Orders and Design", "Public use of reason", "Words; Public; Digital; Citizens", "Politics of mediation", and "Rethinking the Human Condition". The diagram is framed by a large, stylized circular shape with a central vertical line. Four larger text blocks are positioned around the perimeter: "The public sphere in a computational era" (top-left), "Hyperconnectivity" (top-right), "Complexity, responsibility, and governance" (bottom-left), and "Identity, selfhood, and attention" (bottom-right).

Vielen Dank für Ihre Aufmerksamkeit!

Judith Simon

Technologies in Practice Group
IT University Copenhagen
Rued Langgaards Vej 7
DK-2300 København
jusi@itu.dk

Institute of Philosophy
University of Vienna
Universitaetsstr. 7
A-1010 Vienna
judith.simon@univie.ac.at