

Forschungsinfrastruktur in den Sozial-, Verhaltens- und Wirtschaftswissenschaften und die Arbeit des RatSWD

Gert G. Wagner

13. Januar 2011

- Bestandsaufnahme
- weiterführende Strategien
- zukünftige Aufgaben

Bestandsaufnahme

Rückblick

- Wissenschaftliche **Großinstrumente** (für astronomische Messungen) gab es schon in prä-historischen Zeiten; Aufschwung ab 9. Jahrhundert (zuerst im islamischen Raum)
- Grundlagenforschung war **Abfallprodukt** praktischer (hier: astrologischer) Fragestellungen
- Naturwissenschaften haben durch die **Re-Analyse** von Beobachtungsdaten im 17. Jahrhundert große Fortschritte gemacht

Bestandsaufnahme

Ausblick

- Jim Gray, Microsoft
hat das **4. Paradigma** der Wissenschaftsentwicklung
ausgerufen:
 - Nach Empirie, Theorie und Simulation
 - „**Daten getriebene Wissenschaft**“

Bestandsaufnahme

Arbeitsauftrag RatSWD

- Verbesserung der Forschungs-Infrastruktur im Bereich der Sozial-, Verhaltens- und Wirtschaftswissenschaften
 - Zugang zu vorhandenen Daten
 - Ausbau der Zahl der Datenbasen
 - Ausbau der Qualität der Datenbasen
 - Nachwuchsförderung
- Instrumente
 - Clearing und Lobbying
 - Vertretung in Beratungs- und Gutachtergremien
 - Internationale Vertretung

Bestandsaufnahme

Konstituierende Prinzipien RatSWD

- Gleichgewichtige Vertretung von Institutionen und Persönlichkeiten
- Wahl der Persönlichkeiten

Bestandsaufnahme

Sitzungen im Berichtszeitraum (August 2008 bis Ende 2010)

- RatSWD: 11
- FDI Ausschuss: 6
- Workshops: 12
- Sonstige Veranstaltungen: 6
- Infostände: 6

Bestandsaufnahme

Sitzungen im Berichtszeitraum (August 2008 bis Ende 2010)

- Arbeitsgruppen: 15
 - Kriminalstatistik
 - Mortalitätsregister
 - Future Data Access
 - Regionaler Preisvergleich
 - Geodaten und Georeferenzierung

Bestandsaufnahme

Sitze in Beratungsgremien

- Statistischer Beirat: 1 (plus 1)
- Zensuskommission: 2; zeitweise 4
- KII (Kommission für die Informations-Infrastruktur): 1
- Economic and Social Research Council (UK): 1

Bestandsaufnahme

Entwicklung RatSWD (2008/09 bis 2010/11)

- Zum Jahresende 2010 gibt es 15 vom RatSWD akkreditierte Forschungsdaten- und 3 akkreditierte Datenservice-Zentren; ein weiteres FDZ hat den Antrag auf Akkreditierung im Dezember gestellt
- Gründungs-Zentren:
FDZ: Statistisches Bundesamt, Statistische Landesämter, IAB und RV (Deutsche Renten Versicherung)
DSZ: MISSY (GESIS) und IDSC (IZA)

Bestandsaufnahme

Entwicklung RatSWD (2008/09 bis 2010/11)

- Akkreditierungen

FDZ: ALLBUS (GESIS), BIBB, DEAS, Internationale Umfrageprogramme (GESIS), IQB, pairfam, RWI Essen, SHARE, SOEP und Wahlen (GESIS)

DSZ: BO (Betriebs- und Organisationsdaten, Bielefeld)

- Akkreditierungs-Antrag

FDZ: EBDC (LMU-ifo Economics & Business Data Center)

Bestandsaufnahme

Entwicklung RatSWD (2008/09 bis 2010/11)

Zwei Schritte vor und einen zurück:

- Entwicklung der Forschungsdatenzentren im amtlichen Bereich
 - Datenmatching schreitet zügig voran
 - alle Bundesländer machen dauerhaft mit ihren Landesämtern mit
 - es müssen Gebühren erhoben werden
- De-Eskalierung im Streit zwischen DIW und StaBuA

Entwicklung RatSWD (2008/09 bis 2010/11)

- Arbeitsgruppen-Ergebnisse
 - Buchveröffentlichung Kriminalitätsstatistik
 - Abschlussbericht Mortalitätsregister, der dem Bundesgesundheitsministerium übergeben werden wird

Bestandsaufnahme

Entwicklung RatSWD (2008/09 bis 2010/11)

- **Forschungsförderungsanträge**
 - 9 Anträge für das BMBF begutachtet
 - 1 FDZ wurde im Auftrag des BMBF evaluiert

- **Nachwuchsschulungen**
 - laufend innerhalb der FDZ/DSZ
 - in Zusammenarbeit mit Hochschulen: Bielefeld, FU Berlin

- Der RatSWD legt Wert darauf, dass in seinen Veranstaltungen Formate angeboten werden, die junge Forscher systematisch einbeziehen.

Entwicklung RatSWD (2008/09 bis 2010/11)

- Publikationen
 - 10 Newsletter
 - 7 Bücher
 - 14 European Data Watch-Artikel
 - 14 Research Notes

Bestandsaufnahme

Entwicklung RatSWD (2008/09 bis 2010/11)

■ Publikationen

- 122 Working Papers mit insgesamt 2031 Downloads via RePEc und SSRN
(hinzu kommen die Downloads direkt von der RatSWD-Seite)
- 1 Spitzen-WP: „The Weirdest People in the World?“ mit 1456 Downloads
(und Berichten in Nature, Science und „Bild der Wissenschaft“)

Entwicklung der „Governance“ der Forschungs- und Statistik-Infrastruktur (2008/09 bis 2010/11)

- Allianz der deutschen Wissenschaftsorganisationen: Grundsätze zum Umgang mit Forschungsdaten (RatSWD WP Nr. 156)
- DFG Programm zur Förderung der Informationsinfrastruktur und Forschungsdaten
- WR AGen zur Forschungs- und Informations-Infrastruktur
- GWK: KII (Kommission zur Zukunft der Informations-Infrastruktur)
- Enquetekommission „Wachstum, Wohlstand, Lebensqualität“ des Deutschen Bundestages (2011 bis 2013): sie soll faktisch auch Vorschläge zur Governance machen

Bestandsaufnahme

Entwicklung der „Governance“ der Forschungs- und Statistik-Infrastruktur (2008/09 bis 2010/11)

- “Riding the wave. How Europe can gain from the rising tide of scientific data” (Final report of the High Level Expert Group on scientific Data. A submission to the European commission, October 2010)
- OECD Global Science Form: Data and Research Infrastructure for the Social Sciences
- DataCite: The International Data Citation Initiative
- APE 2011: Academic Publishing in Europe

Weiterführende Strategien

Evaluierung durch und Empfehlungen des Wissenschaftsrats (WR) (2009) zum RatSWD

- Plattform- sowie Koordinationsfunktion
- Interessensvertretung der Datennutzerinnen und –nutzer sowie der Datenbereitsteller
- Beratungs- und Begutachtungsfunktionen
- Setzen von Standards sowie Evaluation von FDZen und DSZen
- Gemeinsame Anlaufstelle dieser Zentren
- Engagement in der strategischen Planung der Dateninfrastruktur
- Aktive Rolle in der Entwicklung einer europäischen und internationalen Dateninfrastruktur.

Weiterführende Strategien

Evaluierung durch und Empfehlungen des Wissenschaftsrats (WR) (2009) zum RatSWD:

Der RatSWD sollte nicht oder nur wenig berücksichtigte Datentypen stärker in seine Bemühungen einbeziehen. Dies gilt insbesondere für:

1. qualitative Daten, Geodaten,
2. Daten von Wirtschaftsforschungsinstituten,
3. administrative und privatwirtschaftliche Transaktionsdaten,
4. Daten aus dem Bereich der Gesundheitsforschung,
5. Psychologie und Bildungsforschung

Weiterführende Strategien

Organisatorisch

- Verbreiterung der disziplinären Grundlagen durch Vergrößerung der Zahl der Sitze im RatSWD
- neue FDZen und DSZen werden kommen

Weiterführende Strategien

Strategisch

- Agenda Setting national und international, Workshops, Arbeitsgruppen und Begutachtungen von Projektvorschlägen weitere
- Verbesserung Zusammenarbeit der Amtlichen Statistik, der Ressortforschung, privater Anbieter und der akademischen Wissenschaft
- Stärkung der internationalen Zusammenarbeit

Weiterführende Strategien

Inhaltlich

- Verbesserung der Datenlage zu Organisationen und zu „Kontexten“
- Ausnutzung bestehender Großerhebungen für mehr theoriegeleitete und gezielte Erhebungen („Related Studies“)
- Unterstützung neuartiger Datentypen und Erhebungen (u. a. Geodaten, Biomarker)
- Verbesserung der Qualitätsstandards bei Erhebungen und des Datenmanagements

Weiterführende Strategien

Prozedural

- Verbesserung des Datenschutzes
- Verbesserung der Forschungsethik
- Verbesserung der Rewards für Datenerhebung, Datenarchivierung und Datenzugang

Zukünftige Aufgaben

Entwicklung RatSWD

Persönliche Prognosen

- Inhaltlich

Daten über „natürliche Experimente“ werden immer wichtiger werden.

Deswegen werden langlaufende Längsschnitterhebungen, die georeferenzierbar sind und Daten „hypthesenfrei“ sammeln, um im richtigen Moment eine Veränderung messen zu können, zunehmend Bedeutung gewinnen.

Und Governance-Fragen werden sehr schwierig werden!

Zukünftige Aufgaben

Entwicklung RatSWD

Persönliche Prognosen

- Inhaltlich

Nicht nur Transaktionsdaten (Rendtel 2010) , sondern auch geisteswissenschaftliche Datenbestände werden an Bedeutung gewinnen, da sie quantitativ analysierbar werden.

Property Rights-Fragen werden schwieriger werden!
(Google Scholar)

Zukünftige Aufgaben

Entwicklung RatSWD

Persönliche Prognosen

- Inhaltlich

Netzwerk-Daten werden an Bedeutung gewinnen, da sie quantitativ analysierbar werden.

Property Rights Fragen werden ganz schwierig sein!
(Facebook)

Reading the Social Pages

Understanding and Predicting the Demographics and Behavior of Facebook Users

I will describe how Facebook deals with large data sets, how we can learn and make **predictions** at scale, and how we can use our unique data to gain insights into the ethnic composition, political inclinations, geographic distribution, and **sentiments of our user base.**

Chang is a member of the Data Science team at Facebook. There he explores **Bayesian probabilistic modeling, topic modeling, data mining**, and their application to large-scale systems. Jonathan earned his B.S. in Electrical and Computer Engineering from Caltech in 2003, and inches ever closer to finishing his PhD from Princeton.



Jonathan Chang
Facebook

Segnungen des Datensammelns



Soziales Netzwerk Facebook:
Mit einer Datenbank von
einem Petabyte am Puls der
Zeit. (© dpa)

“Seit Monaten tingelt Jonathan Chang, der in der Datenforschungsabteilung von Facebook arbeitet, mit der Botschaft über wissenschaftliche Kongresse, dass Facebook die größte Datenbank angelegt habe - mehr als ein Petabyte - die der Sozialforschung je zur Verfügung gestanden habe.“

These: *Zusammen mit repräsentativen Surveys wird konventionelle Sozialforschung möglich werden!*

Zukünftige Aufgaben

Entwicklung RatSWD

Persönliche Prognosen

- Organisatorisch

(Fach)Bibliotheken werden als Datenarchivare, die auch Meta Daten zur Verfügung stellen, an Bedeutung gewinnen. Organisationsfragen werden schwierig sein

- Erfreuliche Entwicklung: Zusammenarbeit DSZ Betriebs- und Organisationsdaten Bielefeld und Universitätsbibliothek
- DFG-Projekt zum Ausbau „European Data Watch“: ZBW in Zusammenarbeit mit Berlin, Lüneburg und München

- Sicherung FDZ im Bereich amtlicher Daten !!!
 - Ad-hoc Lösungen und (weitere)
Grundsatzdiskussionen (z.B. über BStG)
 - Novellierung Bundesstatistikgesetz ?
- Verbesserte gemeinsame Informationen und
Zugangsportale zu FDZ und DSZ; Ausbau DOI
 - Mehr Koordination/Vernetzung der Akteure
(Bibliotheken, Archive, Wissenschaft)
- Datenschutz und Forschungsethik

- Sicherstellung von angemessenen Autorenschaften
- systematische Nachwuchsschulung und Förderung
- Mehr Transparenz bei der Direkt-Forschungsförderung durch BMBF
- Permanente AG „Zensus-Analyse und Zensus-Vorbereitung“
- Internationale Aktivitäten (open data, open access)