



Forschungsdatenmanagement Erfahrungen und verallgemeinerte Erkenntnisse aus dem Bereich der Sozial-, Verhaltens- & Wirtschaftsdaten

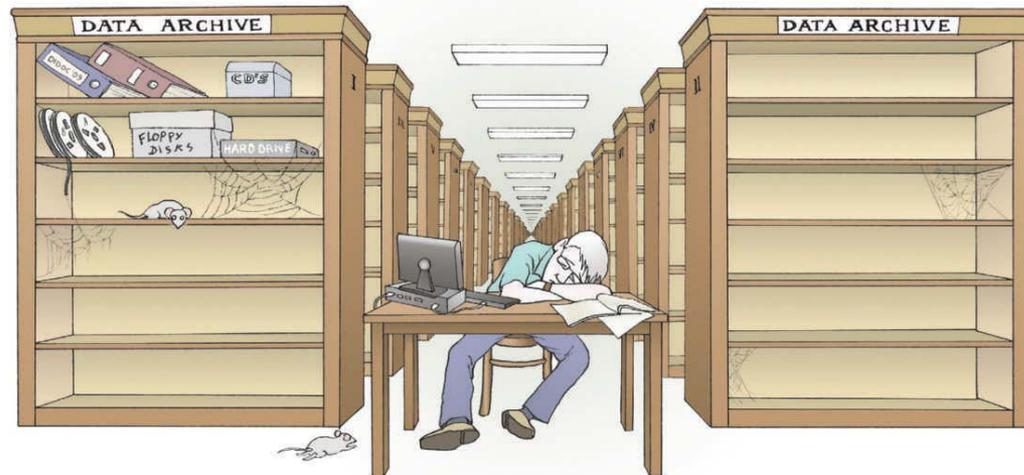
Gert G. Wagner

Interdisziplinäres Symposium

„Umgang mit Forschungsdaten:
Gängige Praxis und neue Anstöße“

Zurückhaltung bei Datenproduktion und -austausch

Data archiving is a good investment
(Piwowar, 2011)



Gliederung

- Hintergrund
- Ausgangslage: Situation in den 90er Jahren
- Erreichtes: Etablierte Forschungsdateninfrastruktur in den Sozial- und Wirtschaftswissenschaften
- Fazit und Wunsch für die Zukunft

Hintergrund

- Daten Ressource und Resultat wissenschaftlicher Forschung
- Mit Zunahme digitaler Daten wächst Notwendigkeit auch die langfristige Nutzbarkeit zu sichern
- Forschungsdaten sind zu einem großen Thema geworden
u.a. wurde von der GWK die KII eingesetzt
[Kommission Zukunft der Informationsinfrastruktur])
- siehe auch http://www.ratswd.de/publ/workingpapers_11.php

Hintergrund

- die Unterscheidung in „small data“ und „big data“ sollte nicht von der Datenmenge her getroffen werden, sondern aufgrund der Governance-Struktur:
 - small data werden „nebenbei“ von reinen Forschungseinrichtungen erzeugt
 - big data werden von Infrastruktureinrichtungen der Forschung erzeugt

Hintergrund

- SoWi-Daten“ bestehen aus quantitativen Daten („Statistiken“) und qualitativen Daten (z. B. Abbildungen oder Texten). Auch Kontextdaten (z. B. über das Wetter) sind quantitative Daten

Grundsätzlich gilt:

- erst Re-Analysen machen aus Daten wirklich Forschungsdaten
- Re-Analysen erfordern im ersten Schritt Archivierung

Hintergrund

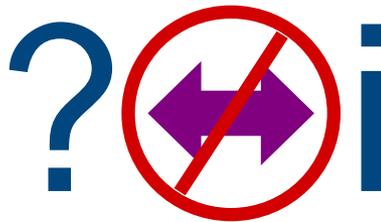
- **Datenmanagement und Data Sharing: Erfahrungen in den Sozial- und Wirtschaftswissenschaften**
Denis Huschka, Claudia Oellers, Notburga Ott und Gert G. Wagner
184/2011
- **Aktuelle Herausforderungen für die wissenschaftliche Informationsinfrastruktur**
Sabine Brünger-Weilandt
RatSWD Working Paper 168/2011
- **Grundsätze zum Umgang mit Forschungsdaten**
Allianz der deutschen Wissenschaftsorganisationen
RatSWD Working Paper 156/2010
- **On the Respective Roles of National Libraries, National Archives and Research Data Centers in the Preservation of and Access to Research Data**
John Pullinger and Gert G. Wagner
RatSWD Working Paper 153/2010
- **Rat für Sozial- und Wirtschaftsdaten (Hrsg.), Building on Progress- Expanding the Research Infrastructure for the Social, Economic and Behavioral Sciences, 2011**



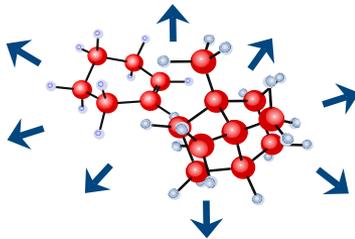
Blick zurück: Mitte 90er Jahre



Zugang zu Daten der **amtlichen** Statistik
nahezu unmöglich



Austausch zwischen amtlicher Statistik
und Wissenschaft kaum vorhanden

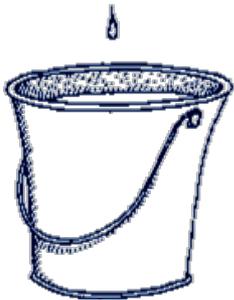


Möglichkeit der effektiven
Anonymisierung von Mikrodaten wurde
als schwierig bewertet

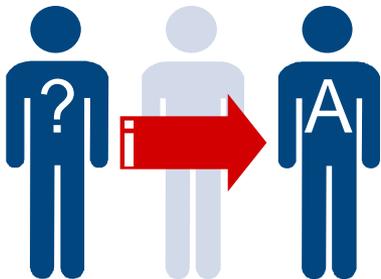
Situation damals...



Datenzugang nur unter
prohibitiven Kosten



Verfügbare Daten ohne hinreichende
Dokumentation



Wissenschaftler mussten sich direkt an
die Verwaltung wenden, kein
nutzerfreundlicher Zugang

Der Wendepunkt...

- **1998:** Memorandum von Hauser, Wagner, Zimmermann „Erfolgsbedingungen empirischer Wirtschaftsforschung und empirisch gestützter wirtschafts- und sozialpolitischer Beratung“
- **1999:** Kommission zur Verbesserung der informationellen Infrastruktur zwischen Statistik und Wissenschaft (berufen durch BMBF)
- **2001:** Veröffentlichung des Berichts mit 35 Empfehlungen, darunter: Einrichtung von Forschungsdaten- und Datenservicezentren, PUF sowie eines Rates für Sozial- und Wirtschaftsdaten- noch im selben Jahr Gründungsausschuss RatSWD
- **2002:** Die ersten FDZ werden gegründet
- **2004:** Gründung RatSWD

In den Sozial- und Wirtschafts- wissenschaften etablierte Forschungsdateninfrastruktur

- **Schaffung einer grundfinanzierten Service-Infrastruktur**
 - Datenservice Zentren
 - Forschungsdaten Zentren
- **Anschubfinanzierung durch BMBF**
- **Finanzierungs-Pluralität**

In den Sozial- und Wirtschafts- wissenschaften etablierte Forschungsdateninfrastruktur

- **Datenservicezentren**
(Forschungsdatenarchive):
 - Statische Datensätze („small data“)
 - Langzeitarchivierung
 - Datenbereitstellung (und Beratung)

In den Sozial- und Wirtschafts- wissenschaften etablierte Forschungsdateninfrastruktur

- **Forschungsdatenzentren**
 - Ständige Veränderungen („lebende Datensätze“) („big data“)
 - Datenarchivierung und -distribution durch Datenproduzenten (einschließlich Amtliche Statistik und andere staatliche Datenproduzenten)
 - Differenzierte Beratung für die Nutzung komplexer Datensätze

Das Erfolgsgeheimnis

- Die Betreiber von erfolgreichen Datenservice- und Forschungsdatenzentren haben institutionelle Grundfinanzierungen, um neben dem Service auch eigene Forschung betreiben zu können (bis zu 50% der Arbeitszeit der Wissenschaftler/innen)
- Der Wissenschaftsrat unterstreicht die Notwendigkeit der Forschungsfähigkeit ausdrücklich !

Herausforderungen

- Immer mehr Daten: wie lassen sich diese nicht nur verfügbar, sondern auch auffindbar machen?
- Kein Metadatenportal (Daten stecken in fachspezifischen Silos)
- Keine koordinierte Strategien der Langzeitarchivierung (Technik, Kosten, Recht (Copyright, Datenschutz), Organisation)
- Kein systematischer Einbezug von Archivierungs- und Zugangsspezialisten: Verlage und Bibliotheken

Herausforderungen

- Für „kleine“ und potentielle Datenproduzenten von „small data“: keine systematischen Incentives und Mittel für Archivierung und Service

(*no* „Credit Where Credit Is Due“- im schlimmsten Falle unterbleibt die Produktion sinnvoller Forschungsdaten)
- Keine Dokumentations- und Zitationsregeln
- Bislang noch nicht einmal eindeutige Identifikatoren für Daten und Datenproduzenten

Fazit und Wunsch für die Zukunft

- Zugang zu Daten in den Sozial- und Wirtschaftswissenschaften trotz schwieriger Ausgangslage und Berücksichtigung Datenschutz sehr gut.
- **Aber** noch zu wenig Standards für Forschungsdatenmanagement im gesamten Data Life Cycle.

Wunsch für die Zukunft

- Anerkennung Produktion von Forschungsdaten als wissenschaftliche Leistung vergleichbar Publikation.
- Interdisziplinärer Austausch über best practice wie im Rahmen dieses Symposiums – vielen Dank an die Veranstalter!



Vielen Dank für Ihre
Aufmerksamkeit!

gwagner@ratswd.de