



Anforderungen von Seiten der Wissenschaft an Archivierung

Workshop „Archivierung sozial- und
wirtschaftswissenschaftlicher Datenbestände“

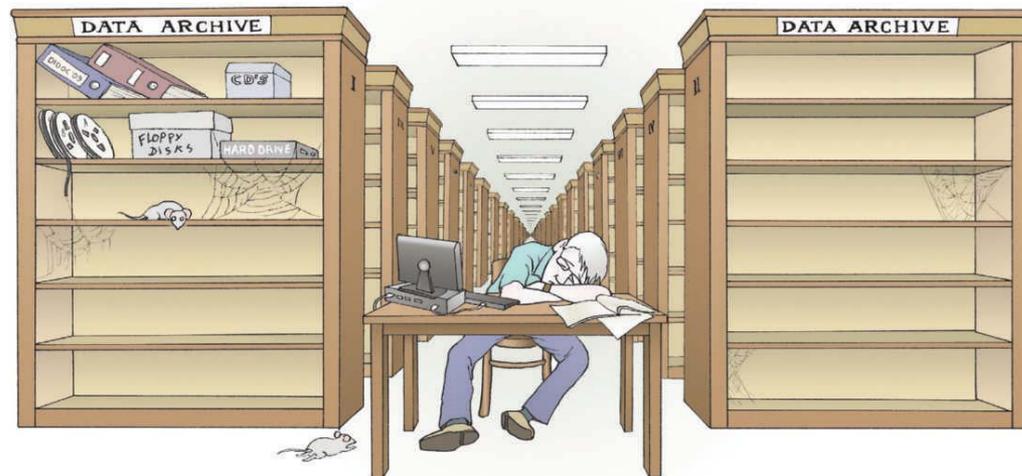
15./16. September 2011,

Deutsche Nationalbibliothek
Frankfurt am Main

Gert G. Wagner

Zurückhaltung beim Datenaustausch

Data archiving is a good investment
(Piwowar, 2011)



Gliederung

- Hintergrund
- Bedarf der Wissenschaft
- Probleme bei der Archivierung von Daten
- Fazit: Voraussetzungen für Data Sharing

Hintergrund

- Vortrag: Beschränkung auf SoWi
- 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

- 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



Bedarf der Wissenschaft: Finden und gefunden werden

Als Nutzer von Daten:

- Gute Dokumentation der Daten
- Transparente und nutzerfreundliche Zugangswege
- Gesicherter, das heisst langfristiger Zugriff

Bedarf der Wissenschaft: Gefunden werden

Als Produzent von Daten

- Sichtbarkeit als Datenautor - Datensätze vgl. einer Publikation durch standardisierte Zitation
- Anerkennung der Datenproduktion als wissenschaftliche Leistung
- Unterstützung beim Archivieren
- Integration des Verfügbarmachens in Datenmanagementpläne und Forschungsförderung

Probleme bei der Archivierung

- Vielzahl an Initiativen, die bislang nicht wirklich koordiniert sind (nestor, RatSWD, DataCite...)

In den Sozial- und Wirtschafts- wissenschaften etablierte Forschungsdateninfrastruktur

- **Datenservicezentren**
(Forschungsdatenarchive):
 - Statische Datensätze
 - Langzeitarchivierung
 - Datenbereitstellung (und Beratung)

In den Sozial- und Wirtschafts- wissenschaften etablierte Forschungsdateninfrastruktur

■ Forschungsdatenzentren

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

Ungelöste Probleme

- Mangelnde Incentives für Datenproduzenten für Archivierung und Service (speziell für „kleine“ Datenproduzenten)

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

Ungelöste Probleme

- Keine koordinierte Strategien der Langzeitarchivierung (Technik, Kosten, Recht (Copyright, Datenschutz), Organisation)
- Kein Metadatenportal (Daten stecken in fachspezifischen Silos)
- Kein systematischer Einbezug von Archivierungs- und Zugangsspezialisten: Verlage und Bibliotheken



Vielen Dank für Ihre
Aufmerksamkeit!