

# gesis

Leibniz-Institut  
für Sozialwissenschaften



## Forschungsdaten kuratieren, archivieren und teilen

### Datenservices des GESIS Datenarchiv für Sozialwissenschaften

*Jonas Recker (jonas.recker@gesis.org)*

# Gliederung

1. GESIS und das Datenarchiv für Sozialwissenschaften (DAS)
2. Herausforderungen in der Archivierung und Bereitstellung sozialwissenschaftlicher Forschungsdaten
3. GESIS Datenservices für verschiedene Zielgruppen

# Was ist GESIS?

- Eine Infrastruktureinrichtung für die Sozialwissenschaften, Schwerpunkt empirische Sozialforschung
- Forschungsbasierter Service
  - ▶ Daten
  - ▶ Integrierte, verlinkte Informationen (Daten, Literatur)
  - ▶ Methodenberatung, Angebote zur Durchführung eigener Studien
- Weiterbildung
  - ▶ Spring / Summer Schools, Workshops
- Mitglied der Leibniz-Gemeinschaft

# Datenarchiv für Sozialwissenschaften

- Gegründet 1960 als Zentralarchiv für empirische Sozialforschung an der Universität zu Köln
  - ▶ seit 2007 Abteilung von GESIS
- Unsere Bestände
  - ▶ Quantitative Daten der empirischen Sozialforschung
  - ▶ Daten zur Erforschung gesellschaftlicher Entwicklungen in nationaler, international vergleichender und historischer Perspektive
  - ▶ > 5.500 nationale und internationale Studien öffentlich zugänglich
  - ▶ 3 Forschungsdatenzentren

# Daten am GESIS Datenarchiv: Beispiele

- ALLBUS – Die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften  
<https://www.gesis.org/allbus/allbus/>
- European Values Study <https://www.gesis.org/evs>
- PAIRFAM – Panel Analysis of Intimate Relationships and Family Dynamics  
<http://www.pairfam.de/>

# Übergeordnetes Ziel unserer Arbeit

- Den langfristigen Zugang zu möglichst hochwertigen Forschungsdaten für Replikation und Nachnutzung ermöglichen
- Herausforderungen
  - ▶ Ethisch-rechtliche Hürden
  - ▶ Erhalt der technischen Zugänglichkeit
  - ▶ Erhalt der Interpretierbarkeit der Daten

# Ethisch-rechtliche Hürden

- Archivierung und Bereitstellung von Daten muss Urheberrechte (z.B. an Instrumenten) sowie Datenschutz bei personenbezogenen Daten beachten
  
- **Ausgewählte Maßnahmen** ✓
  - ▶ Überprüfung der Anonymisierung der Daten bei Eingang
  - ▶ Beratung von Forschenden und Projekten, insb. zu Fragen der informierten Einwilligung, Anonymisierungsstrategien
  - ▶ Kontrollierter und sicherer Zugang zu sensiblen Daten, z.B. über das GESIS Secure Data Center

# Technische Zugänglichkeit erhalten

- Die Nutzung proprietärer Formate (z.B. Statistikprogramme SPSS und Stata) gefährdet die langfristige Zugänglichkeit
- **Ausgewählte Maßnahmen** ✓
  - ▶ Konvertierung in ein Standardformat
  - ▶ „Technology Watch“
  - ▶ Formatmigration



## Interpretierbarkeit erhalten

- Umfassende Informationen über den Erhebungs- und Aufbereitungsprozess notwendig, um Interpretierbarkeit und Nutzbarkeit der Daten zu sichern
- **Ausgewählte Maßnahmen** ✓
  - ▶ Überprüfung der Verständlichkeit und Vollständigkeit der Dokumentation bei Dateneingang
  - ▶ Formale und inhaltliche Erschließung der Daten mit disziplinspezifischen Standards (z.B. [DDI](#), [CESSDA Topic Classification](#))

# Zielgruppen für Datenservices

## Einzelforschende und kleine Projekte:



kleine Fallzahlen,  
kurze Zeiträume,  
geringes  
Datenvolumen

## Kollaborative, langfristig angelegte Projekte:



große Fallzahlen,  
lange Zeiträume,  
aufwändige  
Forschungsdesigns,  
hohes Datenvolumen

## Datenerhebende Organisationen, z.B.



- Außeruniversitäre  
Forschungs-  
einrichtungen,
- Ressortforschung  
des Bundes,
- Universitäten



# Services für Einzelforschende

- Daten für die eigene Forschung suchen und nutzen
  - ▶ [Onlinekatalog „DBK“](#) für Recherche und Download
  - ▶ Nutzung schwach anonymisierter Daten im [Secure Data Center](#)
- Daten teilen: Repository [datorium](#)
  - ▶ Standardisierte Metadaten
  - ▶ [DOI-Registrierung](#)
  - ▶ Eingangskontrolle inkl. Prüfung der Anonymisierung
  - ▶ Bitstream preservation für mindestens 10 Jahre
  - ▶ [Replikationsserver](#) für Daten und Syntaxen zu Journalpublikationen



# Services für kollaborative, langfristige Projekte

- Daten aufbereiten
  - ▶ Added-value Aufbereitung (z.B. Erschließung auf Variablenebene, Harmonisierung von Variablen)
- Daten veröffentlichen
  - ▶ Umfassende Eingangskontrolle (Vollständigkeit, Konsistenz, Anonymisierung, etc.)
  - ▶ Formale und inhaltliche Erschließung unter Nutzung disziplinspezifischer Standards
  - ▶ DOI-Registrierung
- Langzeitarchivierung
  - ▶ Konvertierung in Standardformate, ggf. Formatmigration
  - ▶ Sicherung „für die Ewigkeit“



# Services für Organisationen

- Daten bereitstellen mit SowiDataNet
  - ▶ Standardisierte Metadaten mit DOI-Registrierung
  - ▶ Zweistufiger Kurationsprozess mit umfassender Überprüfung der Daten bei Eingang
  - ▶ Veröffentlichung im GESIS-Webangebot und Einbindung in den eigenen Webauftritt („lokale Vitrine“)
  - ▶ Bitstream Preservation für mindestens 25 Jahre



Entwickelt in Kooperation mit  
ZBW, WZB, DIW Berlin



# Beratung und Schulung

- Forschungsdatenmanagement, u.a.
  - ▶ Einführungsworkshops für Forschende
  - ▶ Beratung, z.B. zu Fragen der informierten Einwilligung und der Anonymisierung
- Langzeitarchivierung von Forschungsdaten, u.a.
  - ▶ Entwicklung von Archivierungsstrategien/Preservation Policies
  - ▶ Vertrauenswürdige Archive (Themenkomplex Zertifizierung)
- Kontakt: [dataservices@gesis.org](mailto:dataservices@gesis.org)

## Für die Zukunft

- Wie können Bibliotheken und Archive bei der Archivierung und Bereitstellung von Daten kooperieren?
- Welche Unterstützungsbedarfe gibt es, die noch nicht gedeckt sind?

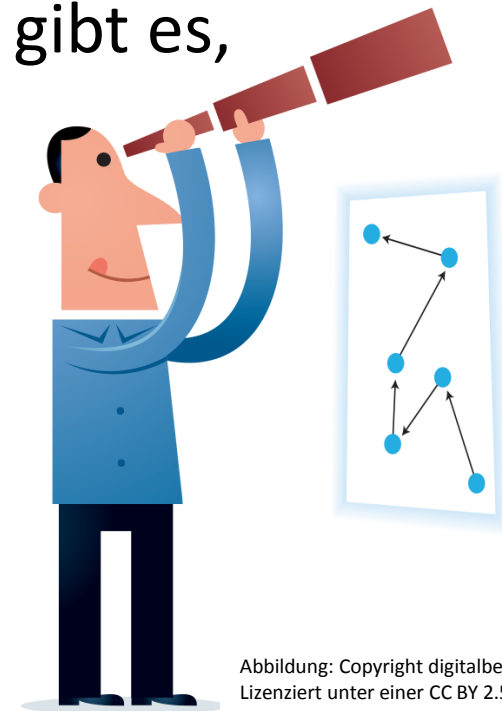


Abbildung: Copyright digitalbevaring.dk.  
Lizenziert unter einer CC BY 2.5 Denmark Lizenz

# Links und Literatur

## GESIS Datenarchiv

- Datenbestandskatalog (DBK): <https://dbk.gesis.org/dbksearch/index.asp?db=d>
- datorium: <https://datorium.gesis.org/xmlui/>
- Sammelrichtlinie:  
[http://www.gesis.org/fileadmin/upload/institut/wiss\\_arbeitsbereiche/datenarchiv\\_analyse/2013-07-08\\_Was\\_wir\\_sammeln.pdf](http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/2013-07-08_Was_wir_sammeln.pdf)
- Secure Data Center: <https://www.gesis.org/sdc>
- SowiDataNet: <https://www.sowidatanet.de>
- Überblicksseite Datenarchivierung: <http://www.gesis.org/angebot/archivieren-und-registrieren/datenarchivierung/>

Mauer, R. 2012. „Das GESIS Datenarchiv für Sozialwissenschaften.“ In: Langzeitarchivierung von Forschungsdaten. Standards und disziplinspezifische Lösungen. Hrsg. Reinhard Altenhöner und Claudia Oellers. S. 197-215. Berlin: Scivero.

[http://ratswd.de/dl/downloads/langzeitarchivierung\\_von\\_forschungsdaten.pdf](http://ratswd.de/dl/downloads/langzeitarchivierung_von_forschungsdaten.pdf)

Recker, J., Zenk-Möltgen, W. und Mauer R. 2017. „Applications of Research Data Management at GESIS Data Archive for the Social Sciences.“ In *Research Data Management - A European Perspective*, hg. von Filip Kruse und Jesper Boserup Thestrup. Berlin [u.a.]: De Gruyter Saur, 119-146.

<https://doi.org/10.1515/9783110365634-008>