

Auswahl von Forschungsdaten

am Beispiel GESIS Datenarchiv

Natascha Schumann & Reiner Mauer
GESIS Datenarchiv für Sozialwissenschaften

nestor-DINI Workshop "Forschungsdaten auswählen und bewerten"
17.11.2015, Universität Duisburg-Essen

Gliederung

- Über das GESIS Datenarchiv
- GESIS Auftrag
- Unterschiedliche Archivierungslevel
- Herausforderungen

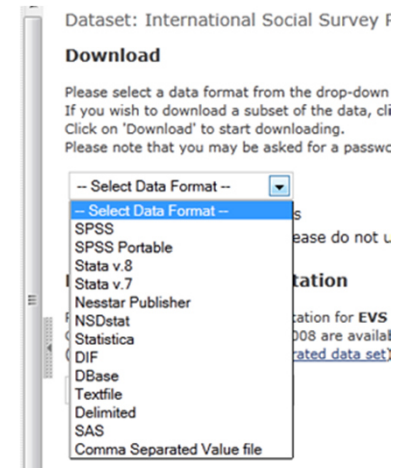
Aufgaben GESIS Datenarchiv

- Langzeitarchivierung von Forschungsdaten: Langzeitverfügbarkeit und Interpretierbarkeit sichern
- Sekundäranalysen, Replikationen, Raum- und Zeitvergleiche ermöglichen
- Zugang zu internationalen Forschungsdaten gewährleisten
- Dienstleister für Primärforscher bei der Sicherung, Dokumentation, Aufwertung und Bereitstellung ihrer Daten

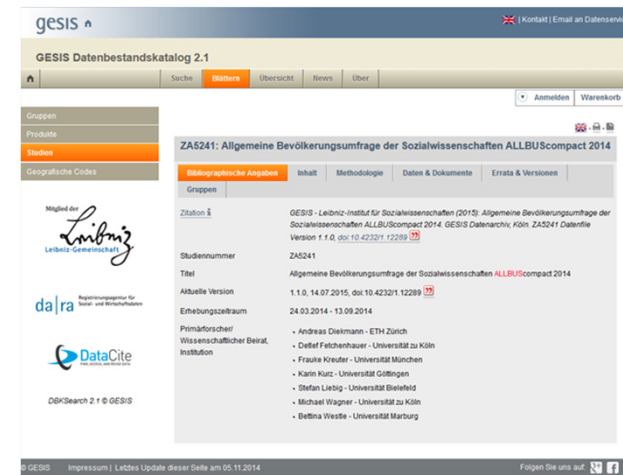


Bestand/Kollektion

- Bestand: mehr als 5500 Studien, insgesamt 620.500 digitale Objekte (2014)
- Nutzung: Ca. 53 000 Downloads (inkl. weitere Vertriebswege) in 2014



- Quantitative Daten
- Gängige Formate
- Datentypen: Umfragedaten, Historische Aggregatdaten...



gesis

„Die Guten ins Töpfchen...“?



„Aschenputtel-Richter“ von Ludwig Richter - wikipedia auf deutsch. Lizenziert unter Gemeinfrei über Wikimedia Commons
- <https://commons.wikimedia.org/wiki/File:Aschenputtel-Richter.JPG#/media/File:Aschenputtel-Richter.JPG>

GESIS-Auftrag

Satzung vom 05.12.2014

§ 2 Zweck

(1) Der Verein dient der Förderung der sozialwissenschaftlichen Forschung. Er erbringt grundlegende, überregional und international bedeutsame forschungsbasierte Dienstleistungen für die Sozialwissenschaften. Er hat die Aufgabe, durch Grundlagenforschung sozialwissenschaftliche Untersuchungsansätze und Forschungsinstrumente zu entwickeln und zu verbessern.

(2) Der Vereinszweck wird insbesondere durch die Erfüllung der folgenden Aufgaben verfolgt:

a) kontinuierliche, interdisziplinäre Forschung und Entwicklung im Zusammenhang mit den in den Buchstaben b) bis g) dieses Absatzes genannten Aufgabenbereichen,

b) Beschreibung und Erklärung gesellschaftlicher Entwicklungen in nationaler, international vergleichender und historischer Perspektive einschließlich der Datenerhebung, statistischen Modellierung und Dauerbeobachtung,

c) Archivierung, Dokumentation und Langzeitsicherung sozialwissenschaftlicher Daten, einschließlich ihrer Erschließung sowie qualitativ hochwertigen Aufbereitung besonders relevanter Daten für Sekundäranalysen,

...

Was wir sammeln

„GESIS hat den Auftrag, Forschungsdaten der empirischen Sozialforschung langfristig zu archivieren und zugänglich zu machen. Die Satzung bestimmt die inhaltliche Ausrichtung des Dienstleistungs- und damit des Datenangebots. Dieses soll der Erforschung „gesellschaftlicher Entwicklungen in nationaler, international vergleichender und historischer Perspektive“ dienen und „international vergleichende Forschung“ unterstützen.“

Merkblatt: Was wir sammeln (Collection Policy)

http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/2013-07-08_Was_wir_sammeln.pdf

Was bedeutet das?

- Sollen wir ALLE sozialwissenschaftlichen Forschungsdaten für IMMER archivieren, die Nutzung garantieren und in GLEICHEM AUSMASS in die Archivierung (Aufbereitung, Dokumentation) investieren?

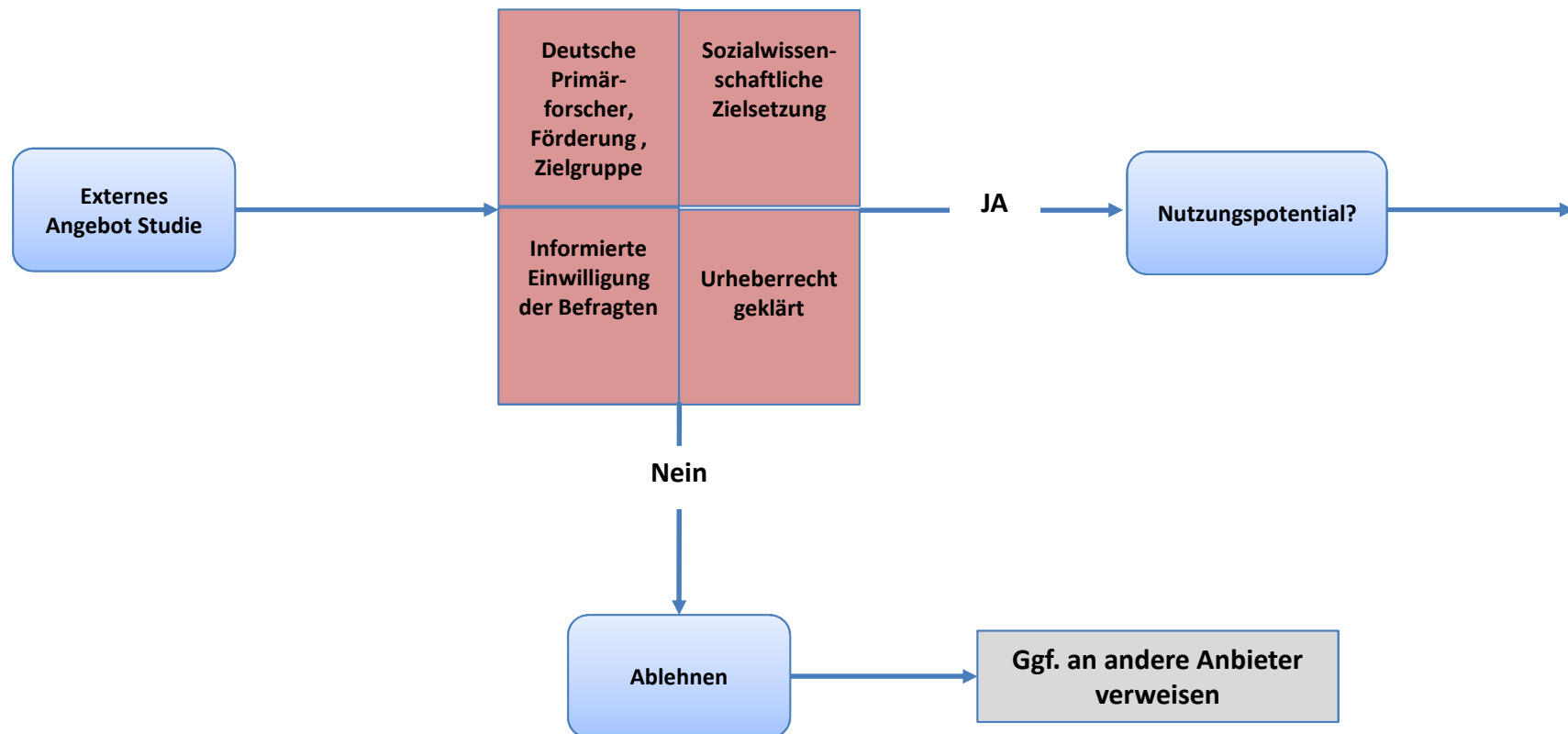
Das bedeutet

- Das können wir nicht:
 - ▶ Zuwachs an Daten → Fehlende Ressourcen
 - ▶ Zu unterschiedliche (Teil-)Disziplinen
 - ▶ Zu viele Formate
- Nicht alle Daten sind gleich, es gibt unterschiedliche Anforderungen im Hinblick auf
 - ▶ Dauer der Archivierung
 - ▶ Qualität und Umfang der Dokumente
 - ▶ Erwartete Nachnutzung

Auswahl

- Mindestanforderungen entscheiden über grundsätzliche Archivierbarkeit/Archivierung.
- Im nächsten Schritt wird versucht, das Nachnutzungspotential zu ermitteln (Entscheidung über das Archivierungslevel).
- Je nach dem werden die Daten unterschiedlich aufbereitet und auch die Art der Archivierung unterscheidet sich.

Mindestanforderungen



Nachnutzungspotenzial

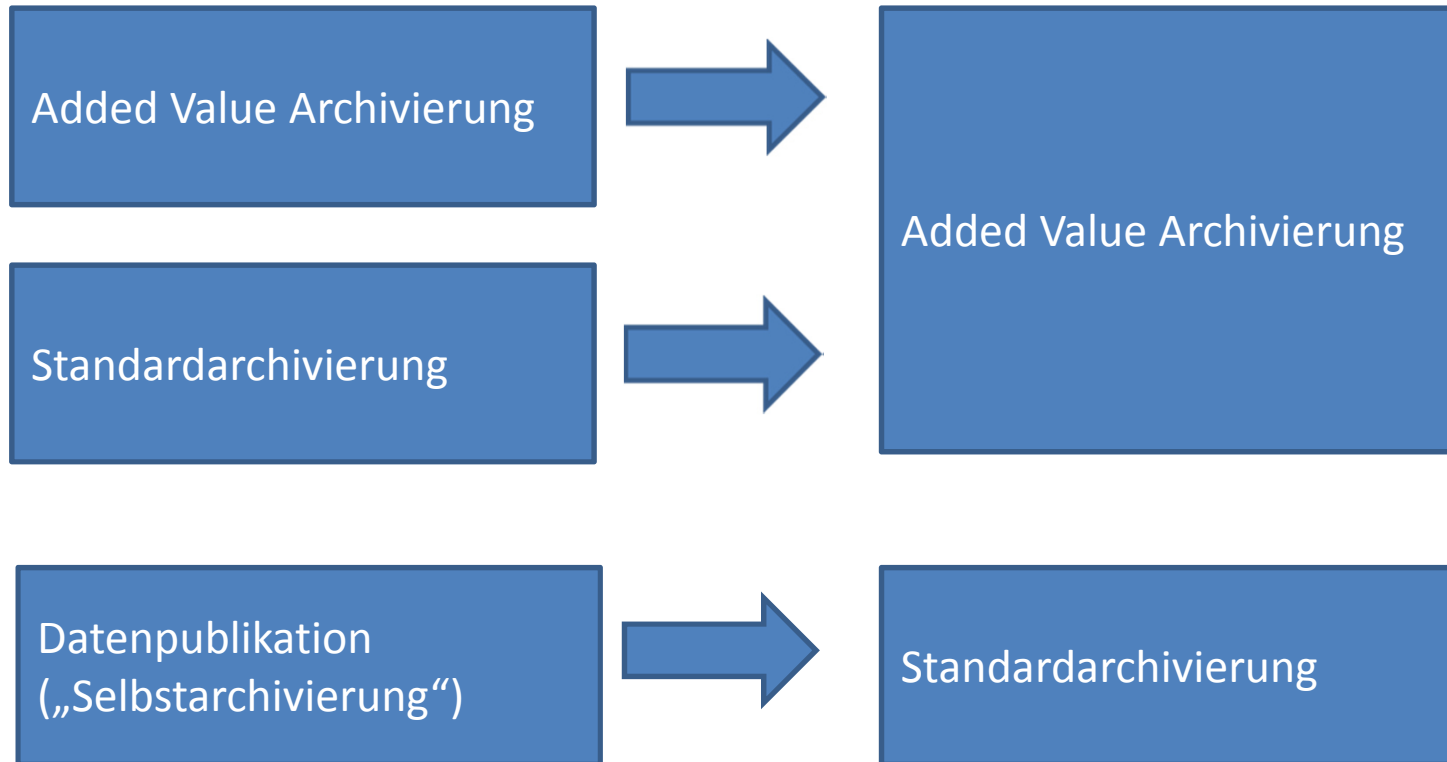
Grundgesamtheit Deutschland	Zufallsstichprobe (falls anwendbar)
Minstdokumentation und Mindestaufbereitung	Datentyp, Datenformat

Füllt Lücken bzw. ergänzt das Angebot	Innovative Erhebungsinstrumente bzw. Methoden
Längsschnittdaten	Kulturvergleichend

Archivierungslevel

derzeit

künftig



gesis

Standardarchivierung

- Minimale Datenprüfung
- DOI Vergabe
- Datennachweis
- Bitstream Preservation



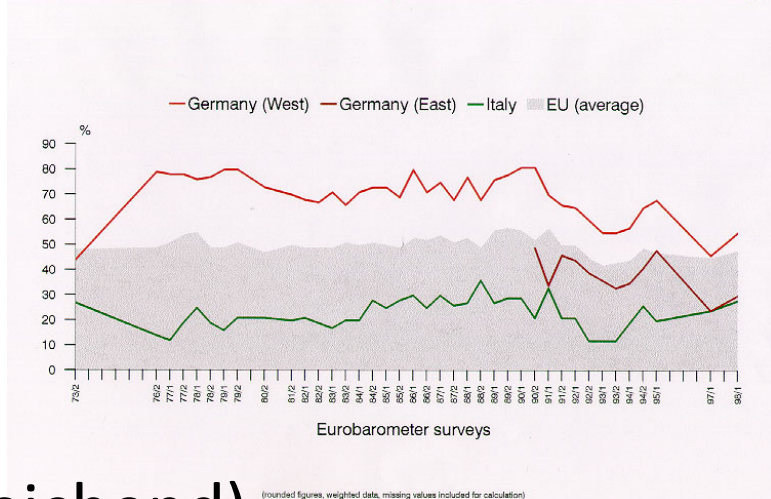
Added Value

- Datenprüfung und Aufbereitung
- Beschreibung auf Studienebene
- Beschreibung auf Variablenebene
- Reports und Berichte
- DOI Vergabe
- Datennachweis
- Langzeitarchivierung
- Bei Bedarf: Bereitstellung im Secure Data Center

Added Value: Daten

- Standardisierung
- Harmonisierung
- Integration / Kumulation
(zeit- und/oder ländervergleichend)
- Ergänzung mit Kontextdaten/
Aggregatdaten

Very or fairly satisfied with the way democracy works in ... (country)



Added Value: Dokumentation

- Umfassende Produktion von strukturierten Metadaten
- Kontextinformationen
- Variablenreports, Datenhandbücher, Methodenberichte,

```

<XML> / <DDI>
<var name="V43">
  <location StartPos="135" width="1"
  <labl> Q3 DEMOCRACY SATISFACTIO
  <qstn ID="Q.3">
    <qstnLit> On the whole,
fairly
    satisfied, not very satisfied
    the way democracy works
    Would you say you are ...?
  </qstn>
  <ivulInstr> READ OUT</ivulInstr>
  <catgry>
    <catValu>1</catValu>
    <txt>Very satisfied</txt>
  </catgry>
  
```



```

<IDNo> ZA4972 </IDNo>
</titlStmnt>
- <rspStmnt>
  <AuthEnty affiliation="European Commission, Directorate General Press and
  Communication, Opinion Polls"> Antonis PAPACOSTAS (Head of unit)
  </AuthEnty>
</rspStmnt>
- <prodStmnt>
  <producer abbr="TNS"> TNS Opinion & Social (original integrated data set
  and documentation) </producer>
  <producer abbr="GESIS"> GESIS - Leibniz Institute for the Social Sciences
  (archive release data set and DDI documentation),
  http://www.gesis.org/ </producer>
  <prodDate date="2010-11-17"> 2010-11-17 </prodDate>
  <prodPlac> Cologne, Germany </prodPlac>
  <software version="4.0.4" date="2011-06-23"> Nesstar Publisher </software>
</prodStmnt>
- <distStmnt>
  <distrbr affiliation="GESIS - Leibniz Institute for the Social Sciences,
  Cologne, Germany" abbr="GESIS" URI="http://www.gesis.org/"> GESIS
  Data Archive for the Social Sciences </distrbr>
</distStmnt>
- <verStmnt>
  <version date="2010-11-17" type="GESIS archive edition">
  Version 3.0.1 (2010-11-17)
  <ExtLink title="Version History and Errata"
  URI="http://info1.gesis.org/dbksearch/sdesc2.asp?
  
```

gesis

Added Value: Beispiele

- Hauptsächlich für *ausgewählte* komparative Studien, die kontinuierlich durchgeführt werden bzw. mehrere Länder umfassen
- Maßgeschneiderte Unterstützung für große Umfrageprogramme
- Häufig ist GESIS bereits an der Datenerhebung beteiligt
- Service u. Forschung dazu i.d.R. durch GESIS FDZs

ALBUS



European *Values* Study



Zusammenfassung

- Mindestanforderungen an alle Daten
- Set an Services für alle Daten
- Differenzierung je nach Art der Studie, nach zu erwartender Nutzung und Ressourcen

Herausforderungen

- Klare Trennung zwischen den einzelnen Levels
- Aufbau der Sammlung sowohl in die „Tiefe“ als auch in die „Breite“
- Neue Datentypen
- Kriterien zur Einschätzung Nachnutzung:
 - ▶ müssen noch weiterentwickelt (operationalisiert) werden
 - ▶ Evaluation bestehender Kriterien
 - ▶ Suche nach weiteren Kriterien

gesis

GESIS Datenarchiv:

<http://www.gesis.org/unser-angebot/archivieren-und-registrieren/datenarchivierung/>

Natascha.Schumann@gesis.org

Reiner.Mauer@gesis.org