

gesis

Leibniz-Institut
für Sozialwissenschaften



Zum Management sozialwissenschaftlicher Forschungsdaten und der Rolle von Archiven

Anja Perry

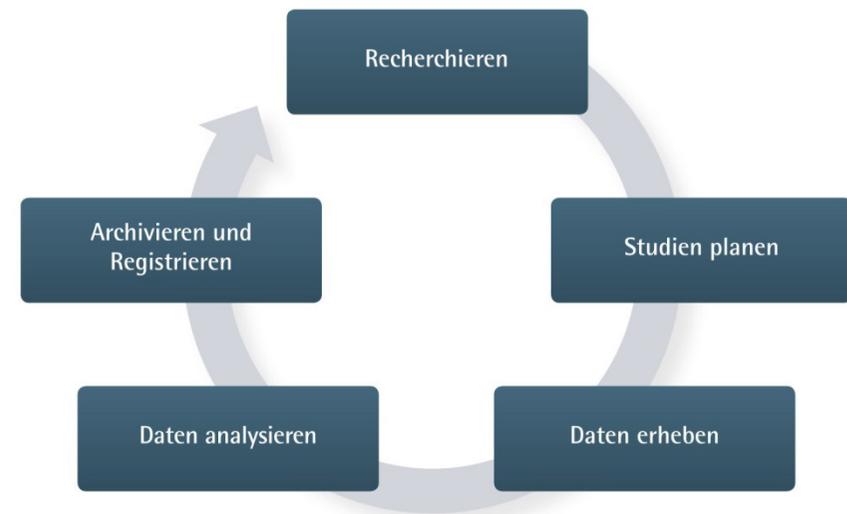
DINI-Jahrestagung, Göttingen, 04.10.2017

Gliederung

1. Was ist GESIS?
2. Was sind sozialwissenschaftliche Daten?
3. Herausforderung in der Archivierung
4. Wie können Forschende unterstützt werden?
5. Ziele des Datenmanagements
6. Offene Wissenschaft

Was ist GESIS?

- Infrastruktureinrichtung für die Sozialwissenschaften
- Schwerpunkt empirische Sozialforschung
- Mitglied der Leibniz-Gemeinschaft
- Forschungsbasierter Service rund um den Forschungsdatenzyklus



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, die Erforschung gesellschaftlicher Entwicklungen in nationaler, international vergleichender und historischer Perspektive ermöglichen
 - > 5.500 nationale und internationale Studien öffentlich zugänglich
 - 3 Forschungsdatenzentren

Sozialwissenschaftliche Daten

The collage features several elements: a spreadsheet with columns labeled v1 through v16 and rows of numerical data; a document with a signature and a stamp; a stack of black floppy disks with a yellow label; and a document with a highlighted green section containing text like "INTEREXPERIENCE = '1') AND" and "DISP_CBA ≠ '1')".

Image: pixabay (CC-0)

Image: pixabay (CC-0)

Herausforderungen in der Archivierung

Interpretierbarkeit der Daten langfristig erhalten

- Umfassende Informationen über den Erhebungs- und Aufbereitungsprozess notwendig, um Interpretierbarkeit und Nutzbarkeit der Daten zu sichern

Technische Zugänglichkeit

- Speichermedien
- proprietäre Softwareformate (z.B. Statistikprogramme SPSS und Stata)

Ethisch-rechtliche Hürden

- Urheberrecht (z.B. Erhebungsinstrumente)
- Datenschutz bei personenbezogenen Daten (Recht auf informationelle Selbstbestimmung)

Interpretierbarkeit der Daten langfristig erhalten

- Wissenschaftliche Daten sollten mindestens 10 Jahre aufbewahrt werden
- Nachnutzung, um öffentliche Gelder effektiv zu nutzen
- Um die Interpretierbarkeit zu erhalten:
 - ▶ Methodenbericht, Feldbericht
 - ▶ Fragebogen
 - ▶ Codebook
 - ▶ Studienbeschreibung, Metadaten
- Datenregistrierung

Herausforderungen in der Archivierung

Interpretierbarkeit der Daten langfristig erhalten

- Umfassende Informationen über den Erhebungs- und Aufbereitungsprozess notwendig, um Interpretierbarkeit und Nutzbarkeit der Daten zu sichern

Technische Zugänglichkeit

- Speichermedien
- proprietäre Softwareformate (z.B. Statistikprogramme SPSS und Stata)

Ethisch-rechtliche Hürden

- Urheberrecht (z.B. Erhebungsinstrumente)
- Datenschutz bei personenbezogenen Daten (Recht auf informationelle Selbstbestimmung)

Datenübergabe – manchmal auch so...



Bild: Pixabay (cc0)

Technische Zugänglichkeit

- Verbesserung der Langzeitarchivierbarkeit
- Softwareumstellungen können in die Zeit der 10-jährigen Aufbewahrungsfrist fallen
- Physische Datenträger können zerstört werden
- Dreifaches Back-Up
 1. Am gleichen Standort
 2. Am anderen Standort
 3. Getrennt vom Netzwerk

Herausforderungen in der Archivierung

Interpretierbarkeit der Daten langfristig erhalten

- Umfassende Informationen über den Erhebungs- und Aufbereitungsprozess notwendig, um Interpretierbarkeit und Nutzbarkeit der Daten zu sichern

Technische Zugänglichkeit

- Speichermedien
- proprietäre Softwareformate (z.B. Statistikprogramme SPSS und Stata)

Ethisch-rechtliche Hürden

- Urheberrecht (z.B. Erhebungsinstrumente)
- Datenschutz bei personenbezogenen Daten (Recht auf informationelle Selbstbestimmung)

Wem gehört das Bild?



Ethisch-rechtliche Hürden

- Wem gehören die Daten?
 - ▶ Dem Forschenden?
 - ▶ Dem Arbeitgeber?
 - ▶ Dem Auftraggeber?
- Welche Daten dürfen erhoben und veröffentlicht werden?
 - ▶ Informierte Einwilligung
 - ▶ Anonymisierung

Ziele des Forschungsdatenmanagements

1. Qualitätssicherung

lesbare, authentische
Daten; Vermeidung von
Datenverlusten

2. Replizierbarkeit

Nachvollziehbarkeit;
Replikation der
Forschungsergebnisse
ermöglichen; Nachweis
der Validität

3. Nachnutzbarkeit

Nutzung der Daten in
neuen Kontexten;
Beantwortung neuer
Forschungsfragen

Bild: CC-0 (Pixabay)

Was machen wir noch?

- Beratung zur Datenarchivierung
- Workshops zu Forschungsdatenmanagement
 - ▶ im Rahmen von CESSDA
 - ▶ hauseigene Kurse
- Datenbereitstellung und Beratung für die wissenschaftliche Nachnutzung
- Bereitstellung sensibler Daten im Secure-Data-Center
- ...

Wir unterstützen die offene Wissenschaft



Vielen Dank für Ihre Aufmerksamkeit

gesis

Leibniz-Institut
für Sozialwissenschaften

Mitglied der
Leibniz
Leibniz-Gemeinschaft

Kontakt

Anja Perry

anja.perry@gesis.org

Tel: +49 221 47694-464