

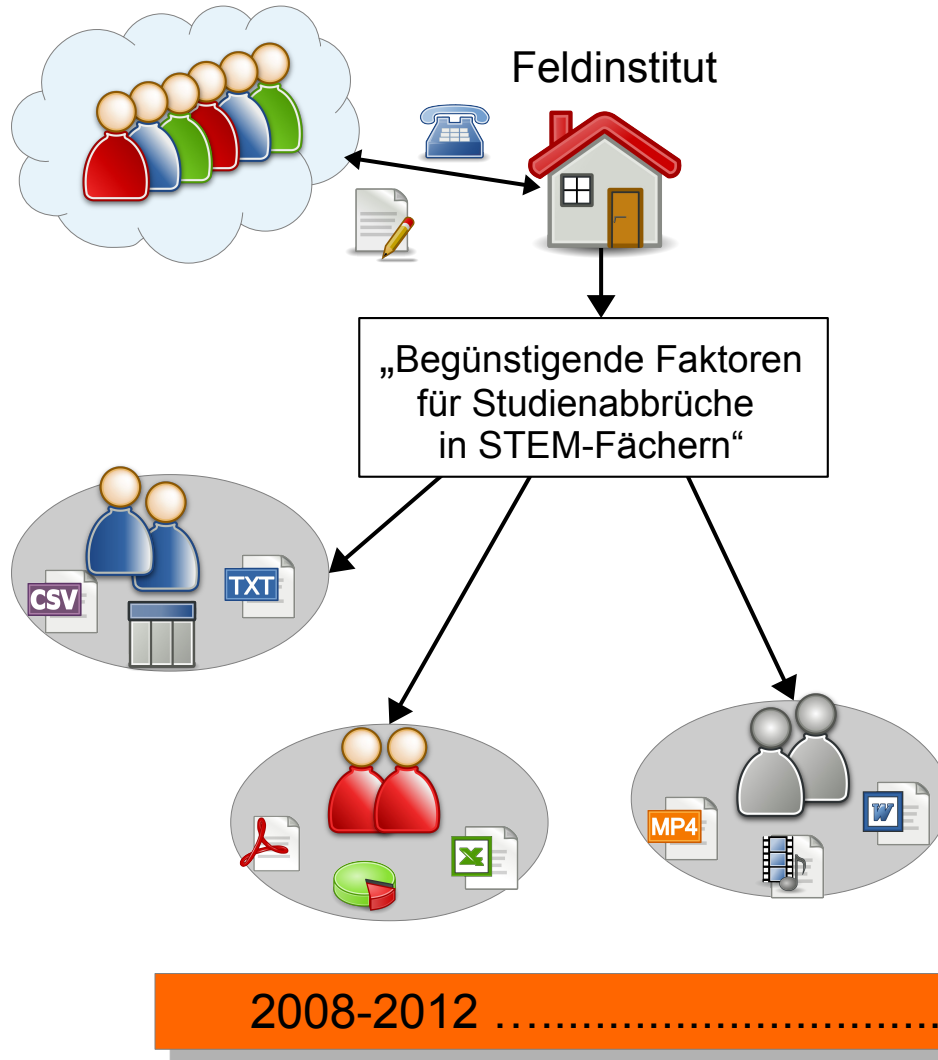
Höher, schneller, weiter? Datenmanagement_(planung) mit Augenmaß!

Astrid Recker | Sebastian Netscher | Jessica Trixa
 GESIS – Datenarchiv für Sozialwissenschaften
 archive.training@gesis.org | @CESSDAtraining

*2. Workshop der DINI/nestor-AG Forschungsdaten:
 „Datenmanagementpläne und ihre Bedeutung im Forschungsdatenmanagement“
 Berlin, 23. März 2015*

Sozialwissenschaftliche Forschungsdaten

- Herausforderungen:
 - Menschliche Teilnehmer_innen = ethische und rechtliche Einschränkungen;
 - Messungen in der Regel nicht oder nur mit großem Aufwand wiederholbar;
 - Umfassende Kontextinformationen notwendig, um Interpretierbarkeit zu erhalten.



Fallbeispiel (fiktiv)

- Nationales Verbundprojekt;
- quantitative und qualitative Daten in mehreren Wellen;
- Werkpakete an verschiedenen Standorten;
- Datenerhebung: Feldinstitut (verwendet eigene Standard-Einwilligungserklärungen);
- Daten werden lokal in den Institutionen gespeichert;
- lokale Regelungen zur Datenaufbereitung und Versionierung;
- zwischen Projektende und Übergabe der Daten an ein Archiv vergehen mehrere Jahre.

Archiv



Probleme im Fallbeispiel

- Informationen über die Feldphase / den Interviewprozess fehlen;
- Vom Feldinstitut verwendete Einwilligungserklärung schließt eine Weitergabe der Daten an Dritte außerhalb der Forschergruppe aus;
- Kodierung und Datenaufbereitung erfolgten nicht einheitlich an allen Standorten;
- keine konsistente Versionierung;
- Uneinheitliche Verwendung von Software und Formaten bei Daten, die zusammengespielt werden müssten (z.B. STATA vs SPSS; ODF vs. DOCX);
- Qualitative Daten wurden nicht anonymisiert/pseudonymisiert;
- Zwischen Projektende und Einreichung beim Archiv sind mehrere Jahre vergangen – Kontextinformationen nur noch mit hohem Aufwand reproduzierbar, wenn überhaupt.

Ziele des Datenmanagements

1. Qualitätssicherung

lesbare, authentische
Daten; Vermeidung von
Datenverlusten

2. Replizierbarkeit

Verständlichkeit;
Reproduzierbarkeit der
Forschungsergebnisse
ermöglichen, Nachweis
der Validität

3. Nachnutzbarkeit

Nutzung der Daten in
neuen Kontexten,
Beantwortung neuer
Forschungsfragen

1. Qualitätssicherung

2. Replizierbarkeit

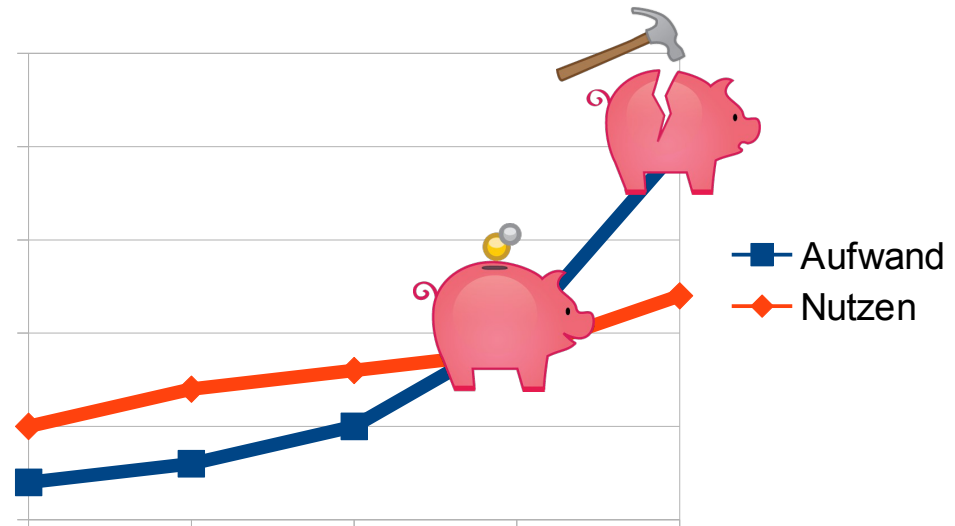
3. Nachnutzbarkeit

Ziele und Maßnahmen des Datenmanagements

Backup-/Sicherheitskonzept (Projektlaufzeit)	Backup-/Sicherheitskonzept (10 Jahre nach Projektende)	Planung der Übergabe an Archiv oder Repository
Richtlinien zur Datenerhebung, -aufbereitung, Versionierung, etc.		Beachtung von Standards (z.B. bei der Skalen- oder Variablenbildung)
Minimale Dokumentation, insbesondere zu Sampling, Interviews, Variablen- und Codelabel	Strukturierte/unstrukturierte Metadaten zur Dokumentation des gesamten Forschungsprozesses (Feldphase, Bereinigung, Analyse, etc.)	Detailliertere Dokumentation (z.B. bis auf Variablenebene; zusätzliche Kontextinformationen)
Klärung ethischer/rechtlicher Fragen zur Datenerhebung, -speicherung und Weitergabe in der Forschergruppe (informierte Einwilligung)	Klärung ethischer/rechtlicher Fragen zur Speicherung und Weitergabe an Dritte	Klärung ethischer/rechtlicher Fragen, <ul style="list-style-type: none"> • die sich bei einer Verknüpfung von Datensätzen ergeben können (z.B. De-Anonymisierung) • die bei der Übergabe an ein Archiv zu beachten sind
Einsatz adäquater Dateiformate für den Projektzweck	Einsatz von Formaten, die (voraussichtlich) in 10 Jahren noch lesbar sind	Einsatz adäquater Dateiformate für die Nachnutzung

Datenmanagement mit Augenmaß

- „Angemessen“ heißt „so viel wie nötig“ um ein angestrebtes Ziel zu erreichen
- Stehen der Nutzen des angestrebten Ziels und der Aufwand in einem angemessenen Verhältnis?



Datenmanagement mit Augenmaß

Um diesen Anspruch einzulösen brauchen wir



Ökonomische
Modelle



(disziplinspezifische)
Kostenmodelle



Werkzeuge zur
Kostenkalkulation