

Herausforderungen beim Aufbau des FDZ:

Gestaltung von Prozessen für qualitatives Forschungsdatenmanagement in der Arbeits- & Industriesoziologie

Heidemarie Hanekop
eLabour / SOFI Göttingen

Was ist eLabour?

- Zusammenschluss arbeitssoziologischer Institute (SOFI, ISF, ISF Jena, SFS, FDZ-BO), die (ihre) qualitativen Studien für Sekundäranalysen bereitstellen und nutzen wollen,
- mit IT- und Infrastrukturdienstleistern (GWDG, SUB Göttingen, L3S) zum Aufbau des ..
- Zentrums und FDZ für qualitative Forschungsdaten in der Arbeits- und Industriesoziologie (AIS)
- BMBF-gefördert als DH Zentrum (10/2015 -09/2020)
- Zukünftig Verein als Träger von FDZ und Zentrum
- Plattform für qualitative Sekundäranalysen in der AIS

Archivierung qualitativer Forschungsdaten aus der Perspektive von Forschenden als Datenproduzenten und –nutzern

- Qualitative Studiendaten sicher archivieren
- Datenschutz und Ethik bei Weitergabe sicher stellen
- Für wissenschaftliche Sekundäranalysen sinnvoll aufbereiten
- Für Sekundärforschung zugänglich machen, kooperieren, zitiert werden
- Sekundärforschung durchführen

Aufgaben eines qualitativen Zentrums für Archivierung und Sekundäranalysen

- Sicherung der qualitativen Studien von WissenschaftlerInnen
- Prüfen und bearbeiten der Daten im Hinblick auf Datenschutz- & Ethik-Anforderungen für Archivierung & Weitergabe
- Archivieren & Kuratieren der Daten für wiss. Nachnutzung: ordnen, Primärwissen aufnehmen, beschreiben, kontextualisieren
- Nutzungsbedingungen für Datensätze festlegen, Zugang gewähren und kontrollieren
- Methoden und Tools für Sekundärforschung: Suchen, Auswählen, Zusammenstellen, Auswerten, Zitieren

Herausforderungen guter Archivierung von qualitativen Forschungsdaten (der AIS)

Qualitative Studien sind in vielfacher Weise heterogen, gegenstandsbezogen und in Forschungskontexte eingebettet

- Archivierung für Nachnutzung soll dies möglichst erhalten oder abbilden, z.B. durch Metadaten, Primärdokumente

Originaldaten enthalten systematisch und überraschend sensible Informationen über Teilnehmer und Organisation

- Schutz durch Pseudonymisieren oder Entfernen soweit notwendig, Kontrolle verantwortungsvoller, wiss. Nachnutzung

Gute Aufbereitung und Kuration herkömmlich zu aufwendig

- Standardisierte, IT-unterstützte Prozesse für die Aufbereitung und Kuration qualitativer Daten für wiss. Nachnutzung

eLabour Vorgehen für gutes, qualitatives Forschungsdatenmanagement (FDM)

- Flexible, modulare aber standardisierte Prozesse ermöglichen auf heterogene Daten zugeschnittene Aufbereitung mit vertretbarem Aufwand
- Grundlage ist systematische Beschreibung und Analyse der Primärdaten (F-Ansatz, Empirie, sensibles Material, Art)
- Risikoanalyse der Originaldaten auf Ebene der Dokumente identifiziert und bewertet sensible Informationen
- differenzierte Datenschutzmaßnahmen und Nutzungsbedingungen der Sekundärnutzung
- Metadaten und Datenmodell sind prinzipiell offen, unterstützen aber Archivierung AIS-spezifischer Fallstudien

- Dreistufiger Archivierungsprozess qualitativer Forschungsdaten

Sichere Ablage von Studien (Originaldaten):

- Hochladen (Ingest) der Originaldaten durch WissenschaftlerIn
- Speicherung vorliegender Daten im Auftrag Primärforscher (T1)

Weitergabe zur Aufbereitung in eLabour

Aufbereitung und Kuration in eLabour Infrastruktur & Plattform

- Zwischenspeicher & IT-gestützter Workflow (T2)
- durch WissenschaftlerInnen oder im Auftrag durch eLabour

Freigabe für das FDZ

FDZ für für qualitative Forschungsdaten & Sekundäranalysen

- Nachhaltige Archivierung geprüfter & kuratierter Studien der AIS
- Zugang für wissenschaftliche Nachnutzung & Nutzungsverträge
- Plattform für Sekundärforschung mit IT-gestützten Tools

- Modulare Aufbereitungs- und Kurationsschritte

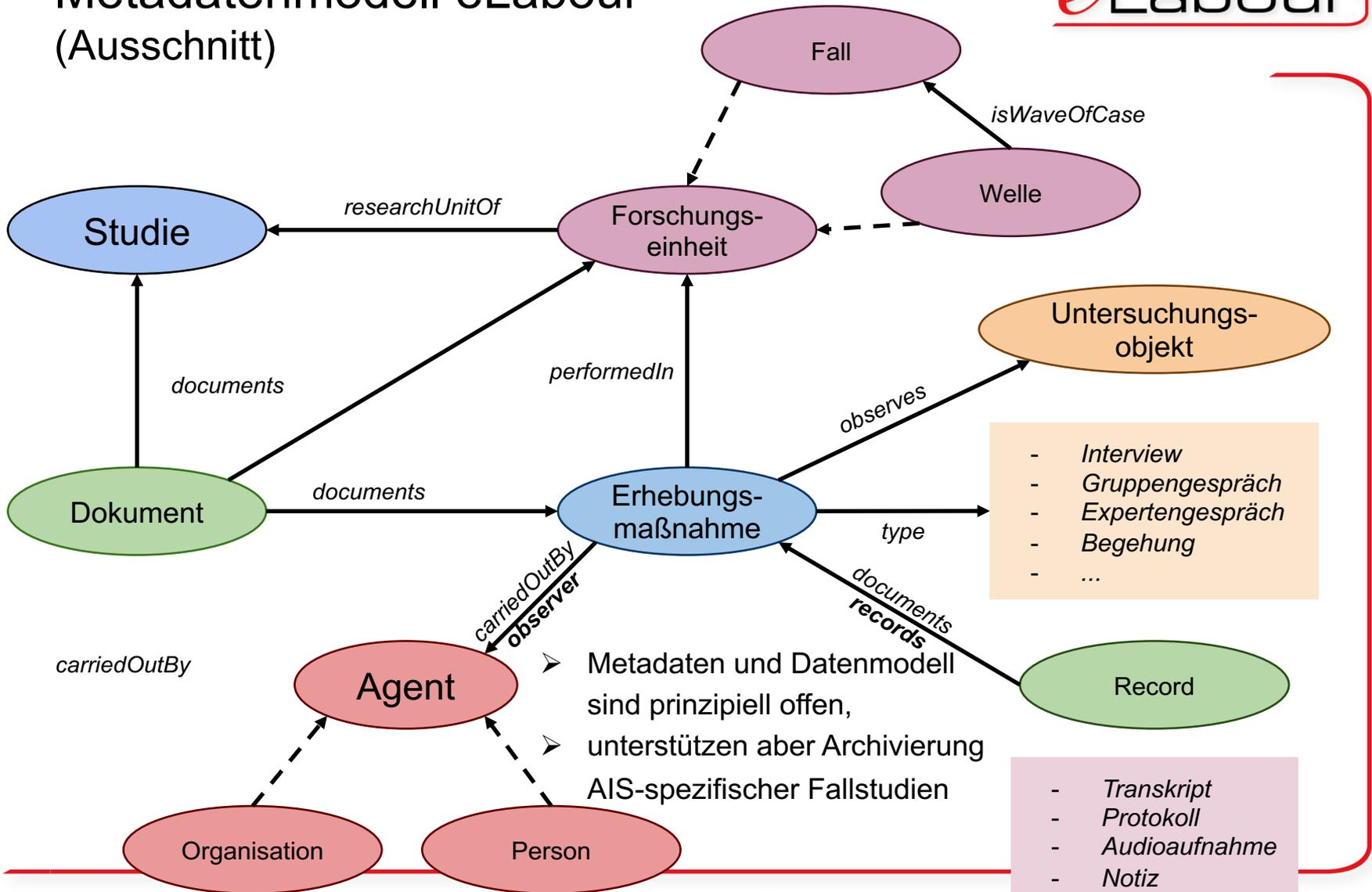
Aufbereitung und Kuration in der eLabour Infrastruktur basiert auf systematischer Beschreibung und Analyse der Primärdaten

Aufbereitung und Kuration in eLabour Infrastruktur & Plattform

- Erschließung und Beschreibung der Primärdaten (F-Ansatz, Empirie, Art und Umfang der Daten, Primärwissen)
- Prüfung und Durchführung von Datenschutzmaßnahmen
- Integration Primärdaten in Datenmodell und Metadaten des FDZ
- Kontextualisierung (Primärstudie, zeithistorisch, themenspezifisch)
- Freigabe für Nutzung mit definierten Zugangs- & Nutzungsbedingungen

FDZ für für qualitative Forschungsdaten & Sekundäranalysen

Metadatenmodell eLabour (Ausschnitt)

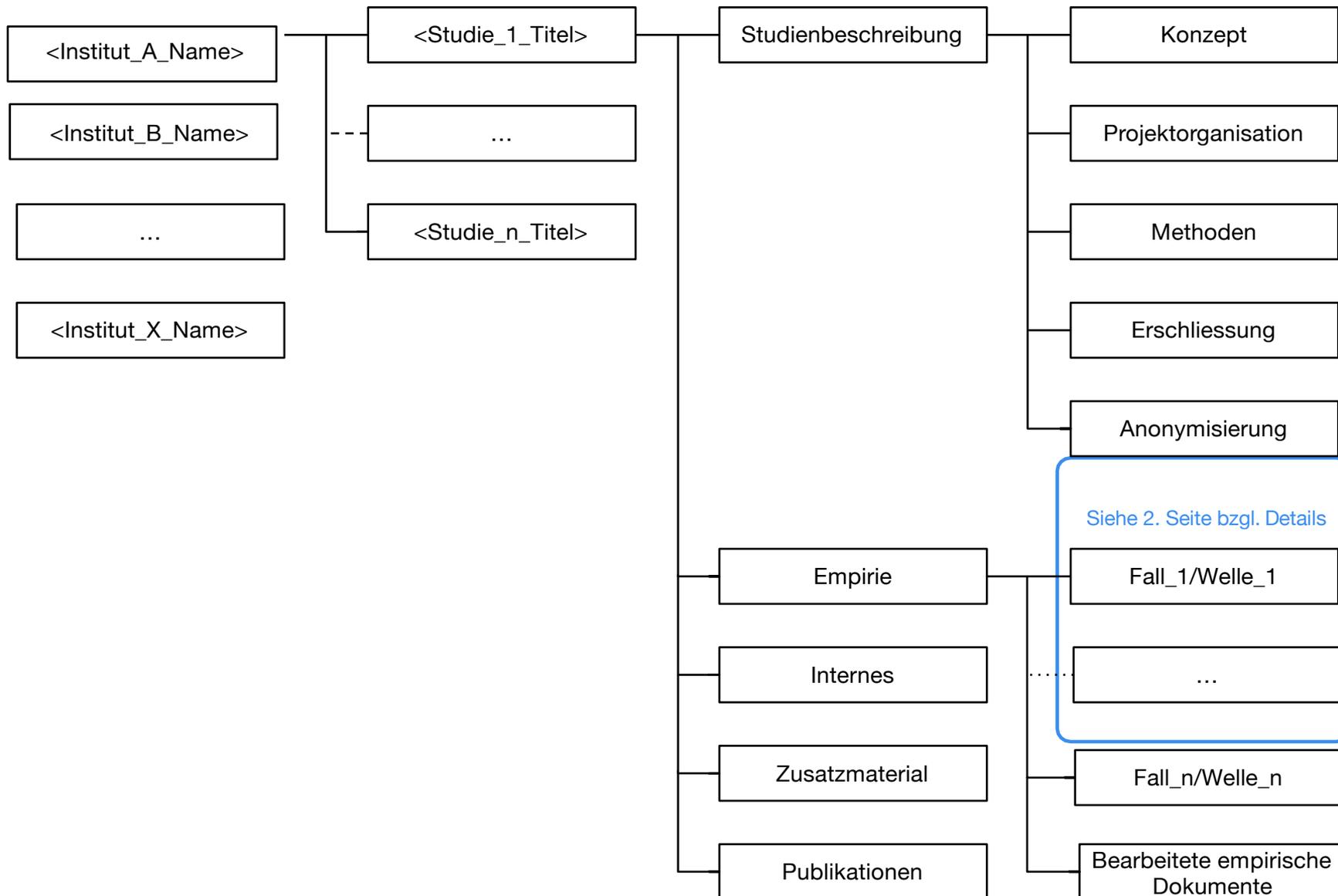


- Interview
- Gruppengespräch
- Expertengespräch
- Begehung
- ...

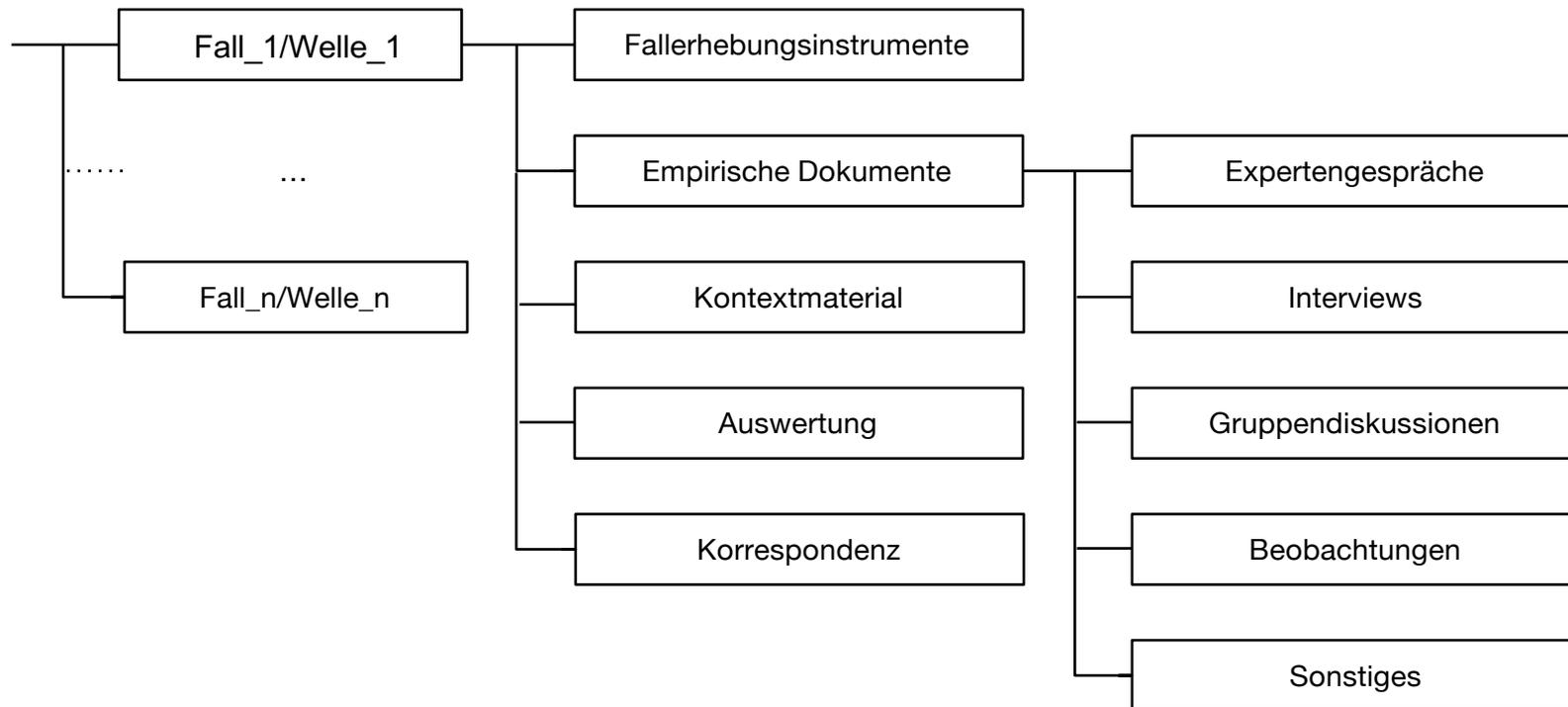
- Transkript
- Protokoll
- Audioaufnahme
- Notiz
- ...

➤ Metadaten und Datenmodell sind prinzipiell offen, unterstützen aber Archivierung AIS-spezifischer Fallstudien

Integration der Primärdaten in Metadatenmodell über Ordnerstruktur (I): die Studienebene



Integration der Primärdaten in Datenmodell über Ordnerstruktur II: die Fallebene



- Namen und direkt identifizierende Informationen pseudonomysiert, getrennt sicher gespeichert (oder gelöscht)
- Weitere Datenschutzmaßnahmen basieren auf konkreter Analyse der Schadensrisiken für Teilnehmer:
 - Risikoanalyse identifiziert und bewertet sensible Informationen in jedem einzelnen Dokument einer Studie & Pseudonymisierung
 - differenzierte Datenschutzmaßnahmen je nach Risikoklasse (TOM)
- Festlegen der Freigabeklasse und Weitergabe an FDZ für Sekundärforschung mit definierten Nutzungsbedingungen

Freigabe- klasse	Schadens- Risiko	Vertragliche Nutzungsbedingungen (pro Studie) generell: Nutzung nur zu wiss. Zwecken, Verbot der Weitergabe, Re-Identifikationsverbot
I. Offener wiss. Zugang	Kein Risiko	Registrierte WissenschaftlerInnen, kein Download möglich (nur remote Lesen)
II. Wissenschaftl. Nachnutzung	Geringes Risiko	Nutzungsvertrag, Download möglich
III. Kontrollierte wissenschaftliche Nachnutzung	Mittleres Risiko	Nutzungsvertrag, Download möglich Datensicherheitskonzept bei Nutzer, Vorlage von Ausschnitten vor Veröffentlichung, Ausschluss von Insidern
IV. Beschränkte wissenschaftliche Nachnutzung	Hohes Risiko	Nutzungsvertrag, Remote Access/kein Download, Datensicherheitskonzept, Schulung der Nutzer, Kontrolle von Ausschnitten vor Veröffentlichung, Ausschluss von Insidern
V. Kein Zugang	Sehr hohes Risiko – keine Nachnutzung	

(differenzierte) Aufwandsabschätzung von Risikoanalyse & Pseudonymisierung

Interviews aus AIS Studien der letzten 10 Jahre:

intensive Betriebsfallstudie, Transkripte:

- mittleres Risiko, ausführliche Erfassung und Extraktion sensibler pers.bez. Info, Pseudonymisierung, ohne Anonymisierung der Organisation
- 60 Min pro ca. 50 S. Interview, Experten- und Gruppengespräche 80 Min

Interviews aus AIS Studien 20-30 Jahre alt:

intensive Betriebsfallstudie, digitalisiert, nicht anonymisiert, nur Protokolle:

- geringes Risiko, Erfassung sensibler pers.bez. Info, umfangreiche Pseudonymisierung mit Schwärzen, Entfernen von sensiblen Notizen
- 30 Min pro ca. 15 S. Beschäftigteninterview oder Expertengespräch

Interviews AIS Studien 30 Jahre alt: digitalisiert, formal anonymisiert, Protokolle

- Kein bzw. geringes Risiko, ggf. Pseudonymisierung mit Schwärzen, Klassifikation sensibler Info,
- 10 Min pro ca. 10 Seiten Beschäftigteninterview oder Expertengespräch

Aufgaben von FDZ und Zentrum eLabour

- Nachhaltige Archivierung geprüfter & kuratierter Studien der AIS
- Infrastruktur und Serviceangebote
- Zugang für wissenschaftliche Nachnutzung
- differenzierte Nutzungsbedingungen für unterschiedliche Freigabeklassen auf der Ebene von Dokumenten
- Nutzungsverträge schließen und Nutzung kontrollieren
- Plattform für Sekundärforschung mit Tools für Suchen, Auswählen, Zusammenstellen, Auswerten, Zitieren qualitativer Daten
- Initiieren und Durchführen von Sekundäranalysen, Methodenentwicklung und Beratung
- Plattform für Zusammenarbeit und Austausch

Herzlichen Dank für Ihre Aufmerksamkeit

www.elabour.de

Heidemarie Hanekop
SOFI Göttingen
heidi.hanekop@uni-goettingen.de