

# Vom Archiv im Keller zur Forschungsinfrastruktur für die Sekundäranalyse: *Eine aufregende Reise*

Ellen Hilf (sfs), Claudia Niederée (L3S),  
Katja Pohlheim (sfs), Stefan Schmunk (SUB),  
Philipp Wieder (GWDG)

**eLabour Konferenz**  
**8. und 9. Februar 2017**  
**Göttingen**



SUB



Niedersächsische Staats- und  
Universitätsbibliothek Göttingen



# Agenda dieser Session

- Einleitung
- Teilprojekt: “Widersprüchliche Integration von Frauen in Dienstleistungsarbeit (WIFE)”
- Prozesse zur Unterstützung der Sekundäranalyse
  - Ingest
  - Datenmanagement und -modellierung
  - Exploration & Suche
  - Sekundäranalyse & Kontextualisierung
- Architektur der eLabour Forschungsinfrastruktur (“Big Picture”)

**Forscher**

Wie lässt sich auf die Daten einfach zugreifen?  
 Was hilft mir bei der Auswahl und Interpretation? Welches Material ist sinnvoll für meine Sekundäranalyse?  
 Wie kann ich mir unbekanntes Material erschließen und welche IT-Werkzeuge helfen mir dabei?

**IT Methodenentwickler**

Was sind nützliche digitale Methoden für eine Sekundäranalyse? Wie lassen sich die Methoden auf diese Formate von Inhalten anwenden? Was sind innovative digitale Ansätze?

**Forschungsdaten-  
infrastruktur****Datenhalter**

Wie kann ich sozialwissenschaftliche Daten dauerhaft und persistent speichern? Was ist der Nutzen für Mich? Wie hoch ist der Aufwand für die Vorbereitung? Wie vermeide ich Risiken?

**Infrastrukturbetreiber**

Wie kann ich die Infrastruktur sicher und effizient betreiben? Wie und mit welchen Werkzeugen kommen die Daten in die Infrastruktur? Was sind die speziellen Anforderungen dieser Daten? Welche Dienste kann ich dauerhaft anbieten und wie kann ich sicherstellen, dass diese weiterentwickelt werden?



# Agenda dieser Session

- Einleitung
- Teilprojekt: “Widersprüchliche Integration von Frauen in Dienstleistungsarbeit (WIFE)”
- Prozesse zur Unterstützung der Sekundäranalyse
  - Ingest
  - Datenmanagement und -modellierung
  - Exploration & Suche
  - Sekundäranalyse & Kontextualisierung
- Architektur der eLabour Forschungsinfrastruktur (“Big Picture”)

## Teilprojekt: Widersprüchliche Integration von Frauen in Dienstleistungsarbeit

### Ziel:

- Befunde zu den Verschränkungen zwischen dem Wandel der Geschlechterverhältnisse und den sektorspezifischen Formen der Reorganisation und Beiträge für Diagnosen zur Entwicklung der gesellschaftlichen Arbeitsteilung
- Untersuchung der Zusammenhänge zwischen den Formen der betrieblichen Organisation von Arbeit und dem beruflichen Selbstverständnis der Geschlechter
- Methode:  
Längsschnittstudie unter Einbezug von vier frühen/früheren sfs-Projekte gendersensibler Arbeitsforschung bzw. arbeitsbezogener Frauenforschung im Einzelhandel und einer Follow up-Erhebung sowie Aufbereitung statistischer Daten zu den Beschäftigungsstrukturen im Einzelhandel

# Arbeit und Dienstleistungsproduktion im Einzelhandel

- Betrachtungszeitraum 1982 – 2017/2018: 35 Jahre
- Zentrale Entwicklungen für Arbeits- und Beschäftigungsbedingungen: Konzentration, Zentralisierung, Deregulierung, Digitalisierung
- Einzelhandel als „Frauenbranche“: früher vergeschlechtlicht, heute nur noch numerisch?
- Wie beziehen sich Personalkonzepte auf Geschlecht, wie auf andere Strukturkategorien?
- Welche Geschlechterarrangements finden sich auf betrieblicher Ebene?
- Welche Persistenzen, Parallelitäten, alte/neue Ungleichheiten finden sich?
- Lassen die Arbeits- und Beschäftigungsbedingungen im Einzelhandel ein berufsfachliches Selbstverständnis (noch) zu?

## Primärstudie „Junge Frauen als Auszubildende und Berufstätige im Warenverkauf – Berufliche Sozialisation, Arbeits- und Lebensperspektiven“

Studie über Zusammenhang von Arbeits- und Lebenssituation, über die Berufsorientierung, das berufliche Selbstverständnis und die Karriereerwartungen junger Frauen (im Vergleich zu jungen Männern) im Einzelhandel

(Förderung durch das Bundesministerium für Jugend, Familie, Frauen und Gesundheit)

1981 – 1985

*Monika Goldmann und Ursula Müller*

*Betriebsfallstudien, Qualitative Interviews und Gruppendiskussionen mit mehr als 140 jungen Verkäuferinnen und Verkäufern in Einzelhandelsunternehmen unterschiedlicher Vertriebsformen; Abschlussbericht und Buchveröffentlichung*

## Weitere Primärstudien:

- „Menschengerechte Arbeitsgestaltung im Einzelhandel unter dem Einfluss neuer Technologien“ (arbeits- und industriesoziologische Studie über die Zentralisierung kaufmännischer und administrativer Funktionen, 1989 – 1992, *Monika Goldmann, Heike Jacobsen, Bärbel Meschkutat, Anita Flacke*)
- „Arbeitsgestaltung im Einzelhandel in den neuen Bundesländern“ (arbeits-, organisations- und wirtschaftssoziologische Untersuchung des sektoralen Transformationsprozesses, 1992 –1995, Heike Jacobsen und Ellen Hilf, Kooperation mit ZAROF e.V. Leipzig)
- „Auswirkungen der neuen Ladenöffnungszeiten auf Arbeitsbedingungen und Beschäftigung im Einzelhandel“ (Repräsentativstudie zur Umsetzung des reformierten Ladenschlussgesetzes, 1998 – 1999, Heike Jacobsen und Ellen Hilf)

## Primärstudien geeignet für sekundäranalytische Untersuchung des Wandels über den Zeitraum

- Entwicklungen der Branche im Fokus, die Arbeit und Beschäftigung massiv beeinflusst haben
- Zentralisierung, Technikeinsatz
- „Osterweiterung“: Deregulierungsschub; gewisse Integration der ostdeutschen „Normalität“ weiblicher Erwerbstätigkeit, zugleich „Backlash“ für die Arbeitsmarktintegration der ostdeutschen Frauen (Teilzeitumutung, Verlust an Status und Beschäftigung zugunsten männlicher Beschäftigter, aufgenötigte Berufswechsel)
- Ausweitung der Ladenöffnungszeiten: Deregulierungsschub, weitere Konzentration
- Methodische Überlegungen: Primärforscherinnen zugleich Sekundärforscherinnen bzw. noch guter Kontakt zu Primärforscherinnen -> Kontextkenntnis

# Exemplarisch: Aufbereiten und sekundäranalytische Annäherung an die „Verkäuferinnen“-Studie

Entscheidung, Verkäuferinnenstudie an den Anfang zu stellen:

- Fragestellung fokussiert auf Arbeitsmarktintegration von jungen Frauen und berufliches Selbstverständnis
- Qualität und Umfang des Primärmaterials: Voll-Transkriptionen von qualitativen Interviews mit mehr als 100 Personen
- Erhellendes Kontextmaterial schon in der Buch-Publikation
- Datenschutz: Material ist „alt genug“, um risikolos eingespeist zu werden, wenn pseudonymisiert/anonymisiert
- „Klassische“ Studie zu Frauenarbeit im Dienstleistungssektor: Material ist es wert, zur Verfügung gestellt zu werden, auch für sozialgeschichtliche Forschung

Kategorien, zu denen im „Überschuss“ des Primärmaterials Aussagen, Beschreibungen und Deutungen zu finden sind

- Geschlechtstypik der Arbeit
- geschlechtersegregierende Personalpolitik
- „Gender status beliefs“
- Geschlechterarrangements
- Statuszuweisungen
- Zuweisungsprozesse bei der Berufswahl
- Berufliche Entwicklungsperspektiven
- Ansprüche auf qualifizierte Berufsarbeit

Kategorien, zu denen im „Überschuss“ des Primärmaterials Aussagen, Beschreibungen und Deutungen zu finden sind

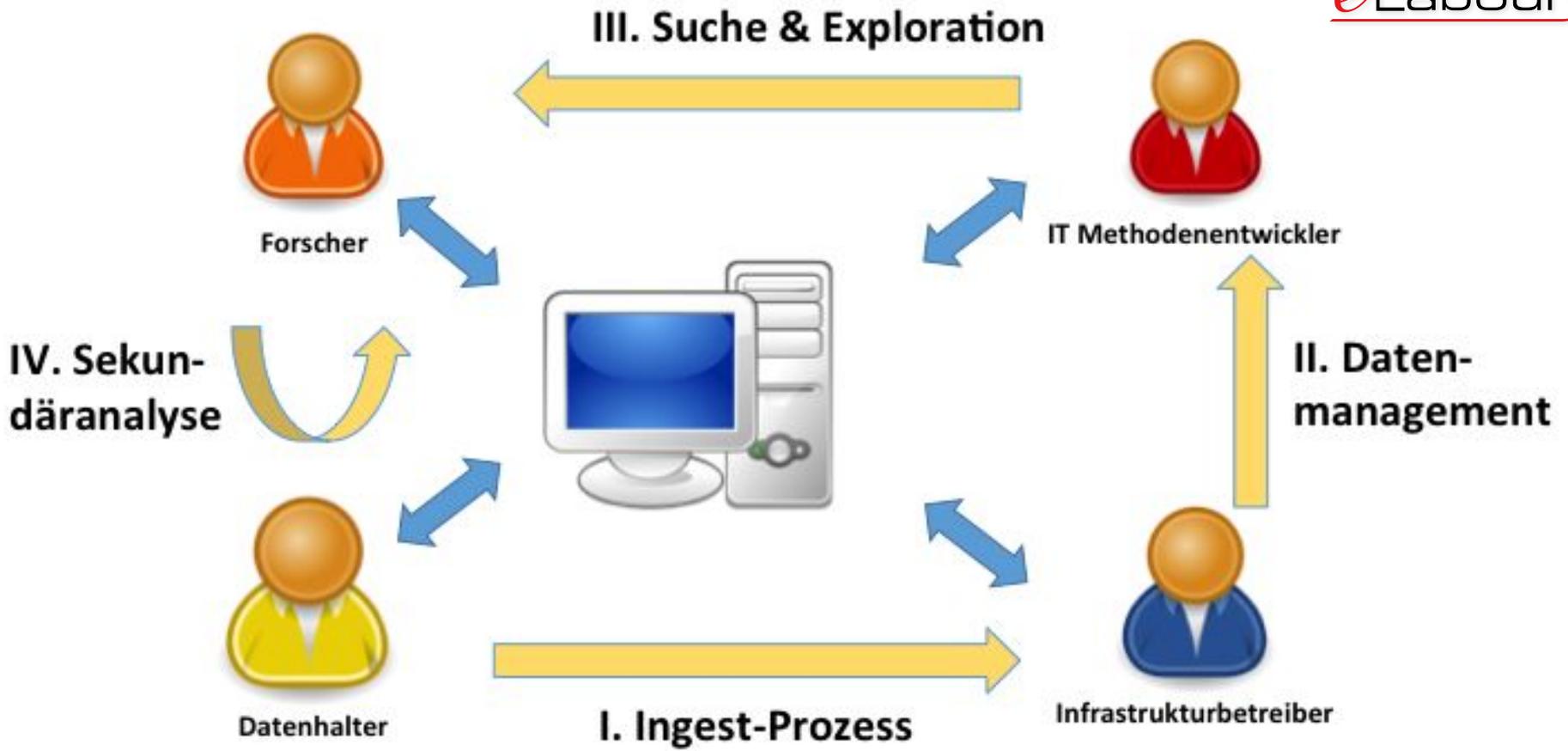
- Interaktion mit Kunden
- Belastungsfaktoren
- Kooperation
- Interessen und Interessenvertretung
- Veränderungen der Verkaufsarbeit (retrospektiv und prospektiv)
- Außerbetrieblicher Lebenszusammenhang, insbesondere Beteiligung an/Verantwortung für Haus- und Versorgungsarbeit; familiäre Arbeitsteilung; private Lebenspläne
- Bewältigung des Vereinbarkeitsdilemmas

# Grenzen der Primärstudie für die Sekundäranalyse

- umfangreiche und stark strukturierte Interviewleitfäden
- an den Leitfäden orientierte dichte Auswertung des Materials durch die Primärforscherinnen
- geringe Anteile narrativer Passagen
- schränkt möglicherweise Eignung des Materials für Sekundäranalysen unter anderen Fragestellungen oder Auswertungsmethoden ein (z.B. für Deutungsmusteranalysen)

# Interdisziplinäre Zusammenarbeit

- Sekundäranalyse bisher “händisch” - zukünftig stärkere IT-Unterstützung
- Für große unbekannte Datenbestände werden digitale Werkzeuge benötigt/  
Datenspeicherung
- Viel Übersetzungsarbeit notwendig - Bsp.: “Fall”  
→ IT braucht gemeinsame Sicht



# Agenda dieser Session

- Einleitung
- Teilprojekt: “Widersprüchliche Integration von Frauen in Dienstleistungsarbeit (WIFE)”
- Prozesse zur Unterstützung der Sekundäranalyse
  - Ingest
  - Datenmanagement und -modellierung
  - Exploration & Suche
  - Sekundäranalyse & Kontextualisierung
- Architektur der eLabour Forschungsinfrastruktur (“Big Picture”)

# Ingest - Lokales Material vorbereiten

- **Ziel:** Material aus arbeitssoziologischen Studien soweit vorzubereiten, dass es in eine Forschungsinfrastruktur integriert werden kann



**Akteur:** Datenhalter

## Umfasst

- Auswahl
- Anonymisierung & Pseudonymisierung
- Aufbereitung (inhaltlich)
- Bereitstellung (Staging - im technischen Sinne)

# Ingest aus soziologischer Perspektive

- Auswahl der Studien:
  - Relevanz
  - Inhalt
  - Struktur



# Ingest aus soziologischer Perspektive



- 6 -

- Ja ich meine ich bin gerne mit Leuten zusammen, ich freundlich und nett sein, das ist klar, man muß auch zu verkaufen. Man kann da nicht bestehen und nicht auch nicht

- also man braucht schon bestimmte Fähigkeiten

- ja sicher

- würden Sie denn auch jungen Männern raten eine Verkaufsmannschaft zu werden, wenn sie die Möglichkeit haben?

- die dann . . . die meisten wollen ja gerne selbst Kaufmann oder so, warum nicht, wenn sie die Möglichkeit haben?

- und wenn Sie das so von heute aus betrachten, würden Sie sich für eine Ausbildung interessieren, oder bleiben, wenn man Verkäuferin lernt?

- eigentlich nicht, brauchte eigentlich nichts geändert zu werden, ich bin mit Kollegen und Kollegen gut auskommen kann, ich das so und so machen oder anders machen, das das dann nicht einem so gefällt, wie die Ware ist

- ach so

- daß man dann mal sagt, anstatt auf Tischen mal in den Regalen

- ja, es gibt ein ganz anderes Bild, wenn ein Kunde auf den Tischen, ist auch nicht gerade so schön, im Regal ist leichter

- verstehe ich es so richtig, daß Sie das damals also das man da so was einbringen kann.

- ja

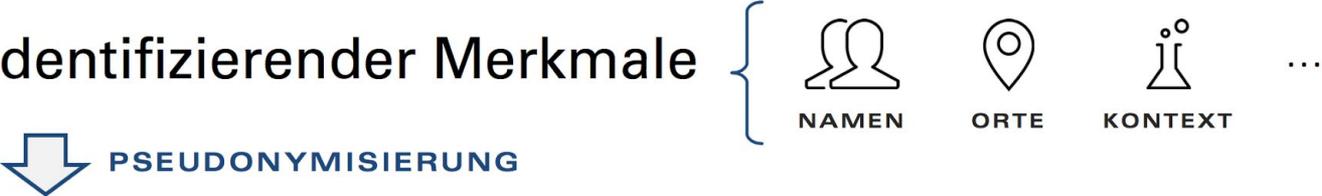
- würden Sie sich denn anders verhalten heute, wenn man den Verkauf machen würden.

- sicher würde ich einiges schon ändern, aber kann man nicht

5 I: Richtig. Gut. @(1)@ Schön. @(2)@ So. Und ich werd erst mal nach'm Gerät  
6 gucken, aber es läuft. Da muss man immer, das muss man (erst mal) im Auge  
7 haben. Gut. Dann würden, würde ich starten. Ich würd s- Sie fragen, was hat  
8 sich denn nach unsrem letzten Besuch, der war zweitausendneun tatsächlich  
9 schon, verändert?  
10  
11 P: Mein Mann ist arbeitslos geworden. (2) Ich bin un-, unplötzlich krank geworden. (6)  
12 Sonst eigentlich (2) würd ich sagen, arbeitslos bin ich nach wie vor, ich kann  
13 gesundheitlich sowieso nicht arbeiten. Und (2) also ich will mal sagen, besser ist es  
14 nicht geworden.  
15  
16 I: Aber, als wir beim letzten Mal da waren, haben Sie noch gearbeitet?  
17  
18 P: Ich nicht. Mein Mann hatte da noch gearbeitet. Ich bin sch-seit vier Jahren schon  
19 arbeitslos.  
20  
21 I: Stimmt. Ihr Mann hat diese Stelle bei der Post gehabt.  
22  
23 P: Ja. Post, Hauptpost saubermachen.  
24  
25 I: Richtig. So. Und Sie waren zu dem Zeitpunkt schon arbeitslos?  
26  
27 P: Ich war da schon arbeitslos.  
28  
29 I: Ja. Und haben Sie (3) in der Zwischenzeit, wenn sie sagen, Sie sind krank  
30 geworden, können Sie mal erzählen, was Sie haben, was passiert ist?  
31  
32 P: Ja. Es, das Einzige, was ich i-, damals Ihnen gesagt hatte, war mit meinen

- Aufbereitung der Studien von z.T.
    - nicht digitalisiert
    - nicht-anonymisiert
  - zu
  - digitalisiert
  - z.T. korrigiert
  - anonymisiert / pseudonymisiert
- Überführung (Ingest/Speicherung) in virtuelle Forschungsumgebung

# Pseudonymisierung

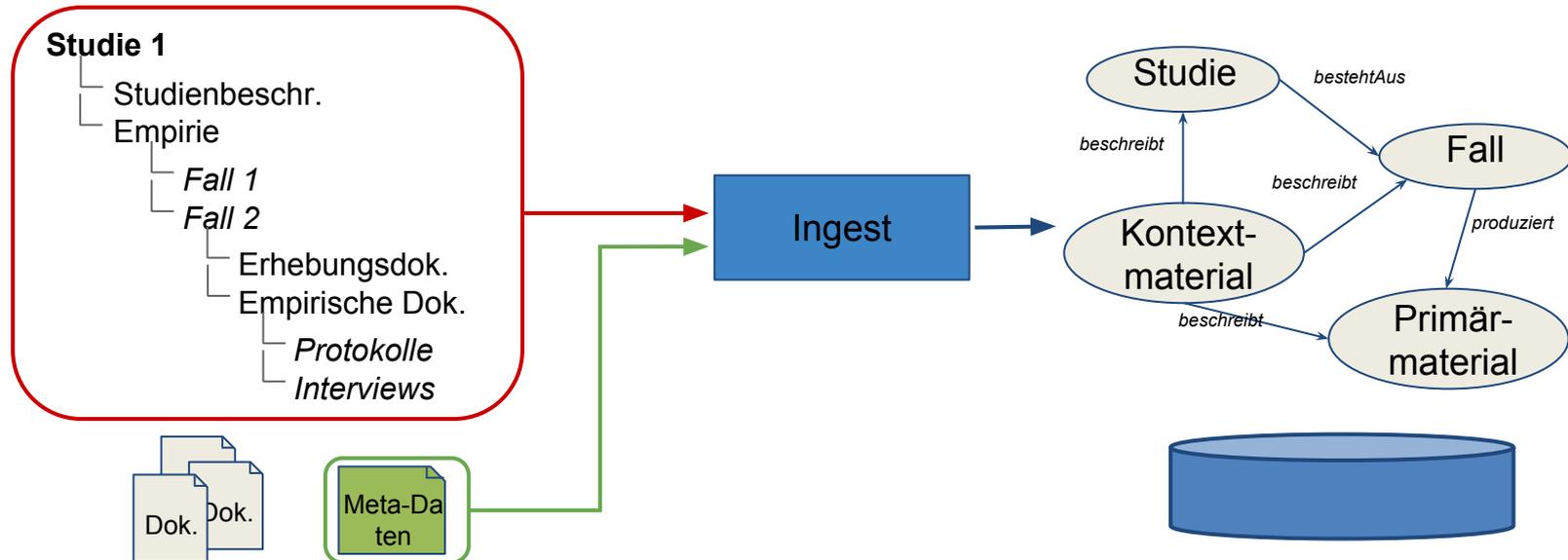
- Entfernen identifizierender Merkmale 

- Tools zur **De-Identifizierung** durch Vergrößerung:  
*Informationsreduktion und Abstrahierung von Daten*
- Flexibel:** Anwendungsbezug und **Kontext** der Daten
  - Varianten für spezielle Forschungsfragen*
  - Auswertbarkeit und Schutz vor Re-Identifikation*

➔ Steuerung des Zugriffs

# Der Ingest-Prozess

- Prozess der Integration von Daten in das Repository der Virtuellen Forschungsumgebung (VFU) mittels Speicherung der Daten in eine Own-Cloud-Lösung



# Der Ingest-Prozess

- Zukünftig mittels Client, der auf den Rechnern lokal bei den WissenschaftlerInnen installiert ist
- Upload in die VFU und halb-automatisierte Anreicherung mit Metadaten

⇒ Daten dort abholen, wo sie bearbeitet werden bzw. liegen

⇒ Anreicherung mit Metadaten durch weitere Tools in eLabour

The screenshot displays the eLabour interface. At the top, it shows 'no changes since last autosave 12:30:18' and a 'show optional metadata' button. The main area is a file tree structure under 'Studie1'. The tree includes folders for 'Zusatzmaterial', 'Empirie', 'Fall3', 'Fall1', 'Fall2', 'Studienbeschreibung', 'Publikationen', and 'Internes'. Under 'Fall1', there is a folder 'Empirische Dokumente' containing 'Expertengespräche', 'Begehungen', and 'Interviews'. The 'Interviews' folder contains two files: 'Interview mit Frau Meier.rtf' (highlighted) and 'Interview mit Herrn Müller.rtf'. Below the file tree, there is a 'view file' button and a metadata form. The form fields are: 'Title' (Interview mit Frau Meier.rtf), 'Creator' (Uwe Sikora), 'Format' (application/rtf), and 'Rights' (empty). A red asterisk indicates that 'Rights' is a mandatory field. At the bottom right, there is a red text prompt: 'Please provide mandatory metadata' and the page number '22'.

# Conclusion Ingest

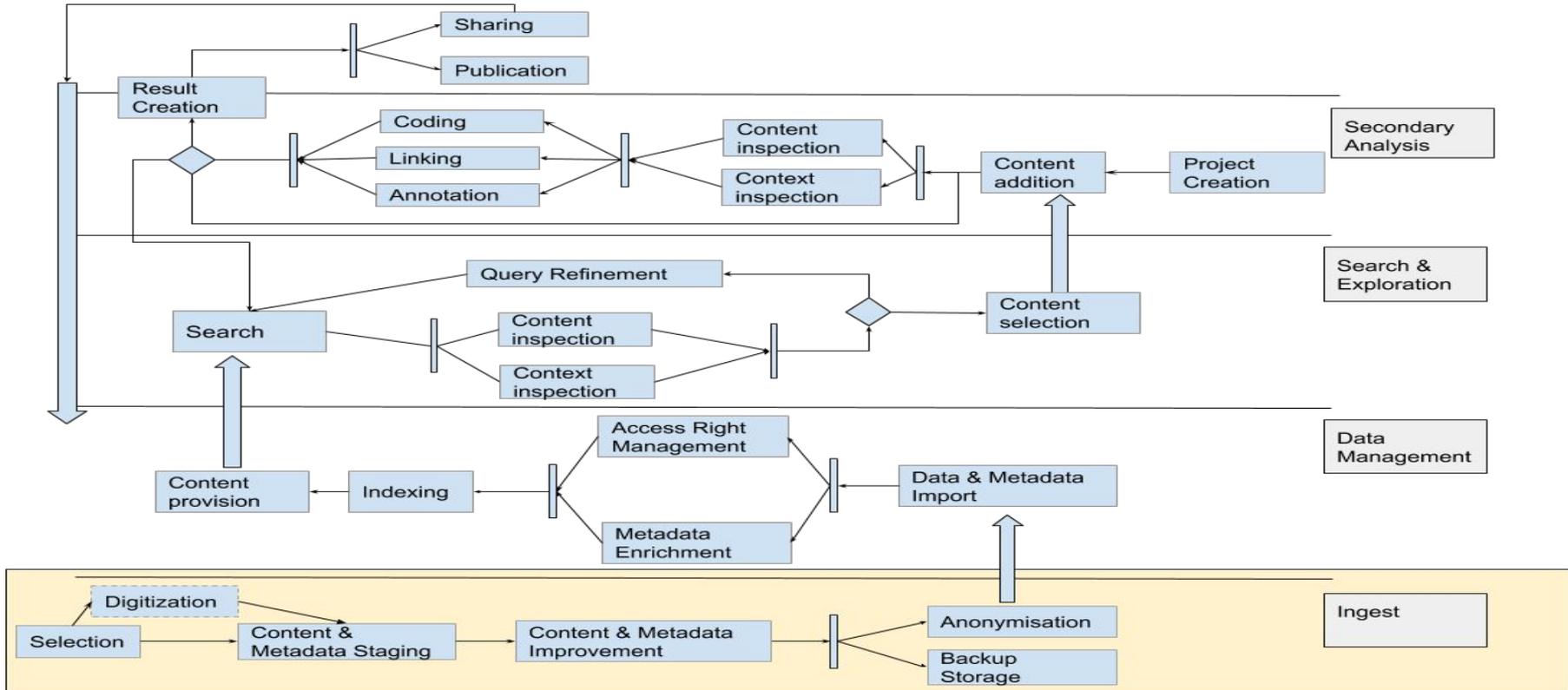
Wo sind die Baustellen?

- Kriterien für Auswahl der Studien?
- Anonymisierung/Pseudonymisierung per Automatismen/Suchfunktionen hinreichend genau?
- Gratwanderung Anonymisierung - Nutzbarkeit der Daten

Was ist gelöst?

- Basisinfrastruktur aufgesetzt
- Speicherumgebung realisiert inkl. Schnittstellen zur Anbindung weiterer Tools
- Ingestprozess konzipiert und Realisierung gerade in der Entwicklung
- Benutzer- und Rollenverwaltung in der Entwicklung

# Kernprozesse Ingest



## Datenmanagement - Material zugreifbar machen und halten

- **Ziel:** Das in die Forschungsinfrastruktur eingebrachte Material in kontrollierter Weise für Forscher zugreifbar zu machen



**Akteur:** Betreiber der Forschungsinfrastruktur

### Umfasst

- (Datenmodellierung)
- Indizierung
- Speicherung
- Zugriffskontrolle

## Datenmanagement - aus soziologischer Perspektive

- Welche Metadaten sind notwendig, um Studien für sekundäranalytische Fragestellungen auszuwählen?
- Wie bekomme ich Zugang zu Studien?



# Eingabemaske Datenerfassung

▶ Werft-Studie  
(BSP-00)

\* ▶ **MockUp-Studie**  
(BSP-01)

▶ Test-Studie  
(BSP-03)

## Studienschließung

**Studientitel:**

**Alternativer Titel:**

+ Weitere Titel hinzufügen

**Forschungsinstitut:**

Name:  Ort:

**Kontakt:**

+ Weitere Kontaktdaten hinzufügen

**Förderer:**

+ Weiteren Förderer zur Auswahl hinzufügen

**Studienlaufzeit:**

Anfang:  Ende:

## Studienschließung

**Studienbeschreibung (Abstract):**

+ Studienbeschreibungen als Dokumente hinzufügen

**Art der Studie:**

+ x

**Forschungsgegenstand:**

+

**Erhebungen:**

1. Beispielerhebung 1

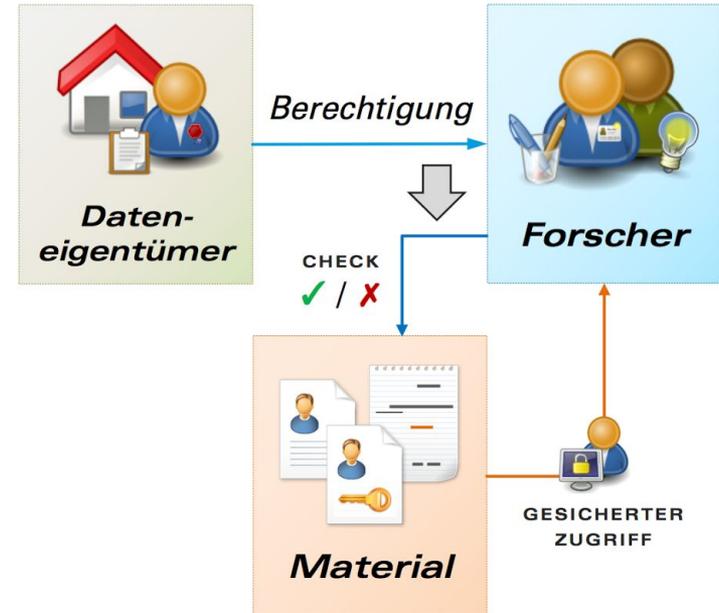
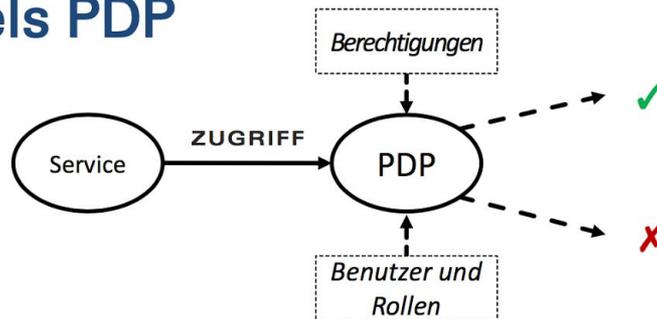
+ Weitere Erhebung hinzufügen

**Veröffentlichungen:**

+ Weitere Veröffentlichungen hinzufügen

# Benutzer- & Rollenverwaltung

- Sicherheit durch nachvollziehbare **Rollen und Berechtigungen**
  - Organisation & Verwaltung der Benutzer und Rollen
  - Zentrale Zugriffskontrolle mittels PDP



# Conclusion Datenmanagement

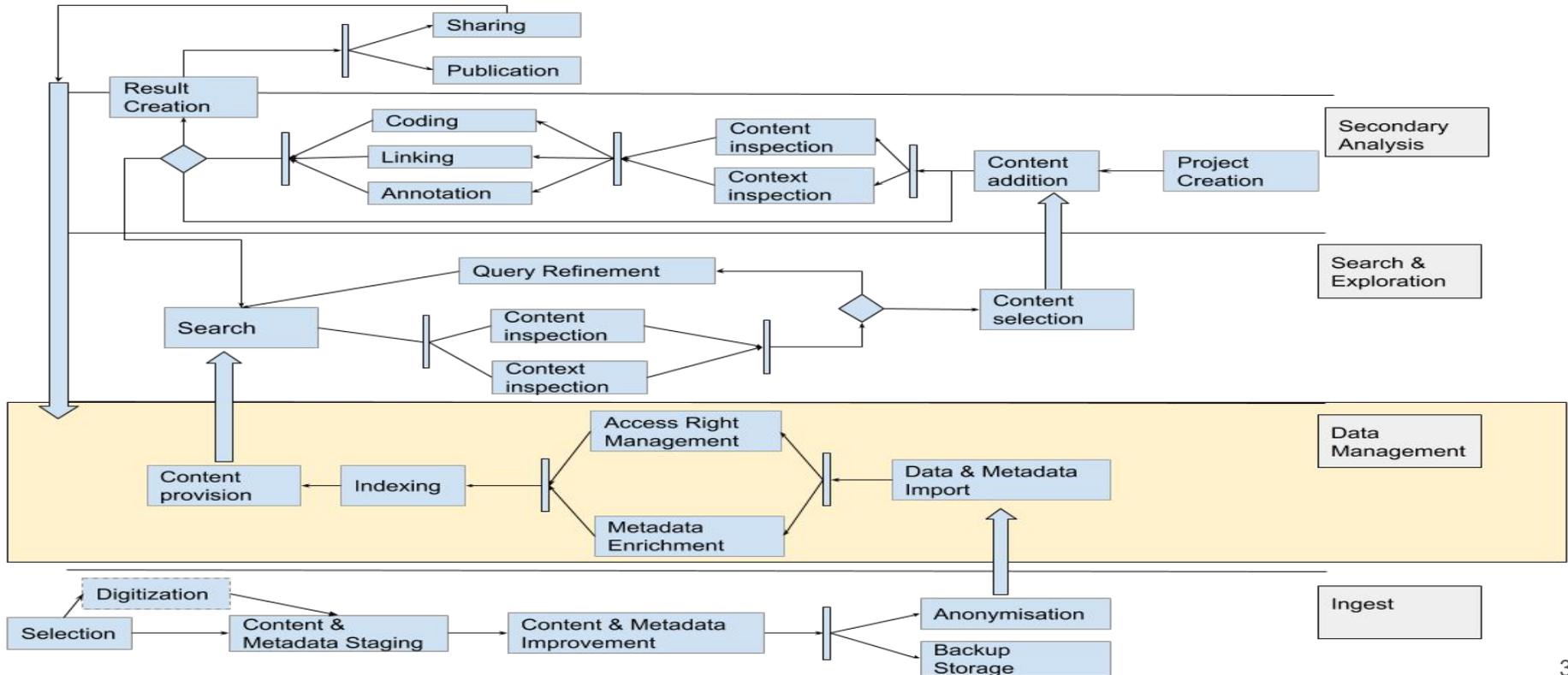
Wo sind die Baustellen?

- Metadateneingabe z.T. sehr aufwändig
- Wieviel können Datenhalter nach Abschluss des Projektes leisten?

Was ist gelöst?

- Datenmodellierung in einer interdisziplinären Metadaten AG umgesetzt
- Modellierung und Programmierung des Interfaces Datenerfassung
- Metadatenstruktur für Dateningest implementiert

# Kernprozesse: Datenmanagement



## Suche und Exploration: Material für Sekundäranalyse finden und auswählen

- **Ziel:** Den Forscher dabei zu unterstützen geeignetes Material für die Sekundäranalyse zu finden, zu verstehen, zu bewerten und auszuwählen.



**Akteur:** Forscher, Betreiber der Forschungsinfrastruktur, Entwickler IT Methoden

### Umfasst

- Suchunterstützung
- Unterstützung zur Inspektion von Suchergebnissen
- Unterstützung für Kontextualisierung

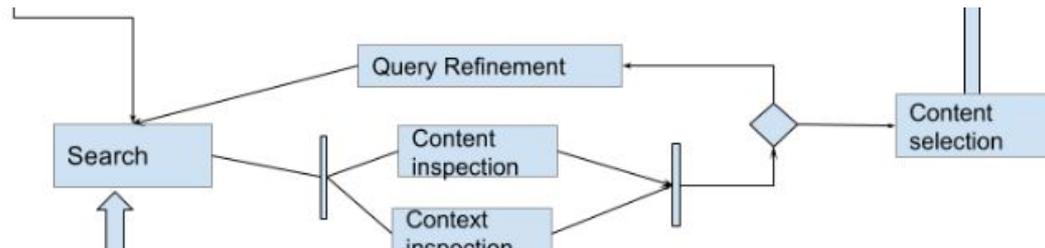
# Suche und Exploration aus soziologischer Sicht

- Wie finde ich relevantes Material für meine Sekundäranalyse?

Auf Ebene von

- Studien
  - Empirischen Dokumenten
  - Absätzen
- 
- Welches Kontextmaterial wird benötigt für die Einordnung des gefundenen Materials?

# Exploration und Suche



- Iterative Suche in den (teilweise) anonymisierten Datenbeständen unterschiedlicher Data Provider
  - Auswahl der Sammlung
  - Beschreibung durch domänenspezifische Facetten
  - Schrittweise Verfeinerung der Suche + Wiederverwendung
  - Unterliegend: Indizierung von Volltext + Metadaten
- Inspektion von Suchergebnissen
  - Darstellung des Inhalts einzelner Dokumente
  - Darstellung der Ergebnisse im Kontext

→ Auswahl von Inhalten für weitere Bearbeitung

# Suche

## Query/Question

Politisches Interesse

## Filters

Zusätzliche Filter

Meine Projekte Filter zurücksetzen

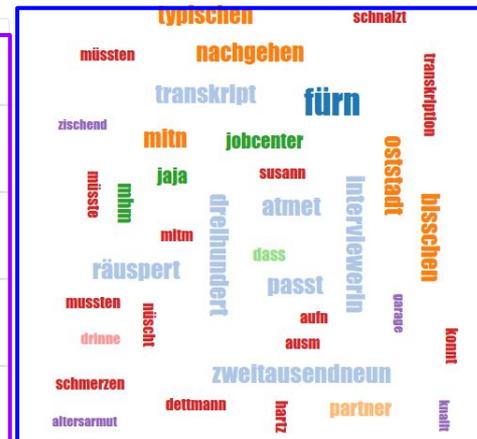
Erscheinungsjahr Institution Studientitel Instrumententyp Forschungsthema Dokumententyp

Eigensinnige Kunden (SFB 580)

7 Suchergebnisse gefunden

## Documents

- 12 Interview Herr Groß**  
Gesellschaft, wo du sagen würdest, die wünsch ich 635 mir? 636 Herr: Das wird aber sehr lang und sehr ausführlich. Sind das alles Fragen, die 637 vorgegeben sind? 638 I: Ich hab so ein bestimmtes politisches Interesse gesehen und dachte, da 639 kann man ja mal gucken. 640 Herr: Ich habe jetzt gerade die entscheidenden Passagen des Manifestes nicht so 641 auf Lager, aber da, eigentlich ist
- 04 Interview Herr Nagel**  
das 517 nem Verlag anzubieten als ne Übersetzung, also, und dann (2) ja, solche Sachen, 518 die nicht, wo also kein Termindruck ist, aber die ich dann so (3) ja, (ist so ne 519 Mischung) aus Interesse, Freizeit und (?) manchmal kann man™s ja auch noch v - 520 werten irgendwie so ökonomisch. @(.)\_@ 521 522 I: Ich frag das, weil wir uns natürlich auch so™n bisschen für (2) ja, die 523
- 06 Interview Herr Koch**  
sollte. (räuspert sich) Bei dem davor, da hab ich erst mal drüber 418 gesprochen (so) und da (hat man), hab ich gesagt, dass mir dieser Job als 419 Produktionsarbeiter hab ich eigentlich kein Interesse, ich möchte was anderes 420 machen. Und dann hab ich gefragt, oder gesagt, was ich gerne machen möchte oder 421 würde. Ja, Grünanlagen, solche Jobs bekommen wir nicht.
- 16 Interview Frau Lampner**  
meisten 125 haben Arbeit oder n paar Studenten sind. Die kommen auch abends nicht zum 126 Grillen oder sonnenabends mittags (?) zu verschiedenen Uhrzeiten gemacht. (2) Aber 127 die haben kein Interesse dran, dann können wir™s auch nicht ändern. Und und und 128 wie gesagt, man hat (ja) die Schlüssel gegenein -, gegenseitig, wenn die mich mal n 129 ganzen Tag h hier nicht erwischen.
- 03 Interview Frau Dreyer**  
Ihre Arbeit sprechen, 396 das ist was, das macht Ihnen Spaß. 397 Frau: Stimmt auch. 398 I: Ja? Ist Ihnen das wichtig, dass Sie was machen, wo Sie drin aufgehen, was 399 auch so bisschen in Ihrem Interesse entspricht? 400 Frau: Ja. Denn ich sage mal, ich hab bei der Firma [ EINZELHANDELSKETTE 1 ], da 401 war das ja dann nachher so, dass der Chef oder wir hatten einen neuen Chef 402
- 11 Interview Frau Bauer**  
informier, was in der Welt los ist oder so, ne. 349 Oder hier am Handy guck ich dann mal (2) interessante Sachen an. Und und les 350 auch gerne, ne. Also das ist wenigstens auch noch geblieben, die Interesse so an 351 ner Außenwelt. Aber, i ch hatte so™n großen, wie ein Burnout. Ich hab nur noch (2) die 352 letzten Monate dann gelegen und geheult. Weil eben in diesen Jahren, wo ich da 353
- 15 Interview Herr Knoll**  
Bildungsgutschein 622 raus gebe, dann würd ich das auch den, (allen) Kindern geben. Ne, es kann ja sein, 623 (2) dass Sie zum Beispiel jetzt, Sie gehen, angenommen, noch in die Schule (4) und 624



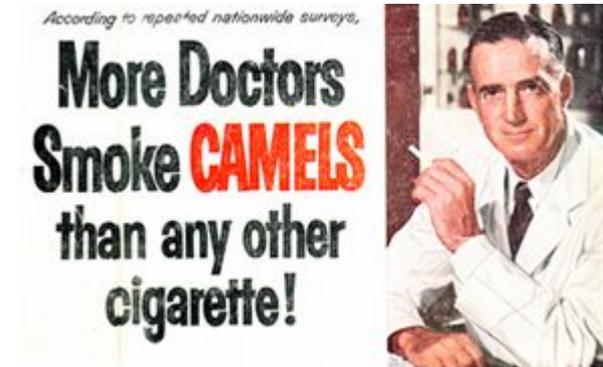
Wird auch oft gesucht

fürn
transkript
räuspert



# Kontextualisierung

Arbeitsdefinition: „Bereitstellung von **zusätzlicher Information**, welche benötigt wird um insbesondere fremde und ältere **Studien adäquat zu interpretieren**“



# Herausforderungen

- Erste Lösungen für allgemeinere Situation auf der Grundlage von Wikipedia

Anwendungsspezifische Herausforderungen:

- Auswahl und Aufbereitung geeigneter Quellen (z.B. Beschäftigungsstatistiken, European Social Survey Data)
- Interdisziplinäre Diskussion: Welcher Kontext ist relevant? (eigene AG)
- Geeignete Methoden zur Kontextualisierung (unterschiedliche Grade der Automatisierung)



# Conclusion Exploration & Suche

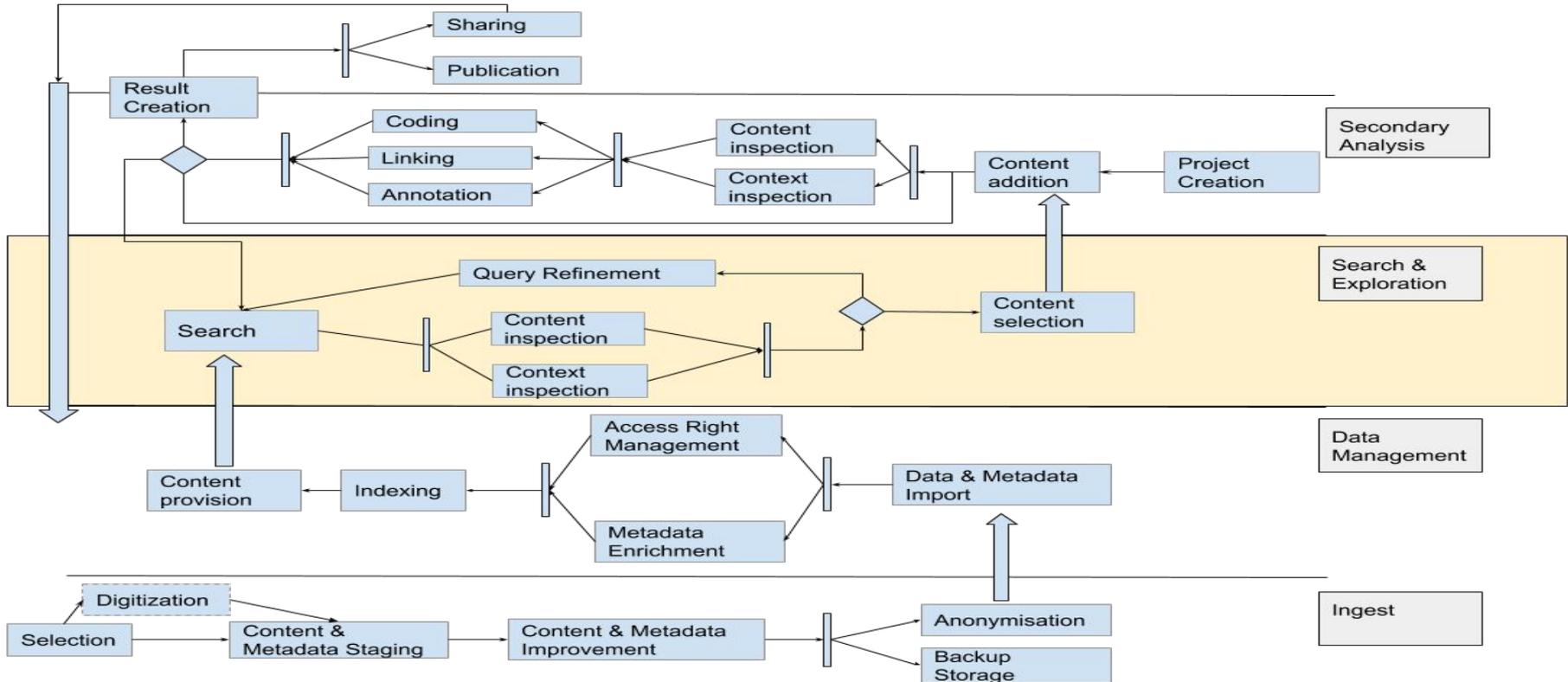
## Wo sind die Baustellen?

- Detaillierte Suche
- Zugang zu zugangsbeschränkten Datenbeständen
  - Andere Forschungsdaten
  - Archive
  - Statistiken etc.

## Was ist gelöst?

- Suche, Facettierte Suche, Darstellung einzelner Dokumente
- Anbindung an Data Management (Kernfunktionalität)
- Noch offen: Wie genau funktioniert die Kontextualisierung?
- Wie gut funktioniert Suche in anonymisierten Dokumenten

# Kernprozesse: Suche und Exploration



## Sekundäranalyse: Neue Erkenntnisse durch Kombination und neue Blickwinkel

- **Ziel:** Den Forscher dabei zu unterstützen eine Sekundäranalyse durchzuführen



**Akteur:** Forscher, Betreiber der Forschungsinfrastruktur, Entwickler IT Methoden

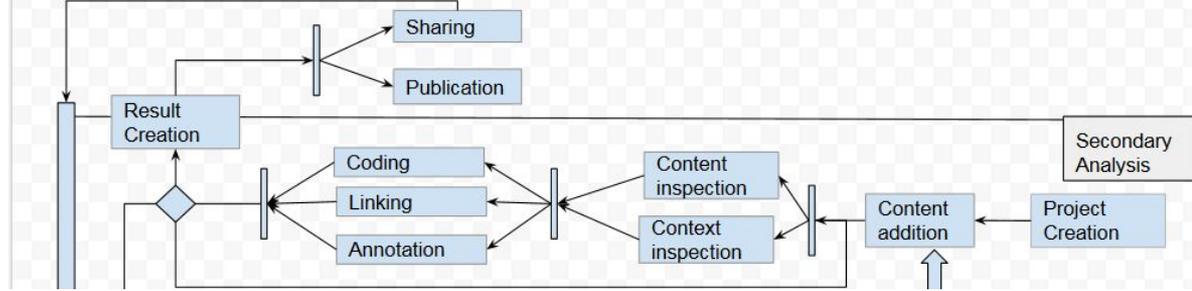
### Umfasst

- Start eines Sekundärprojekts
  - Sammeln von Material
  - Arbeit mit dem Material
  - Export
  - Bereitstellung von Ergebnissen in der Infrastruktur (nicht zwingend)
- (iterativ, siehe Vortrag ISF)

# Sekundäranalyse aus soziologischer Sicht

- Forschungsumgebung für Sekundäranalyse
- Verbindung zu gängigen Analysetools wie Atlas.ti oder MaxQDA
- Kontextualisierung des Materials

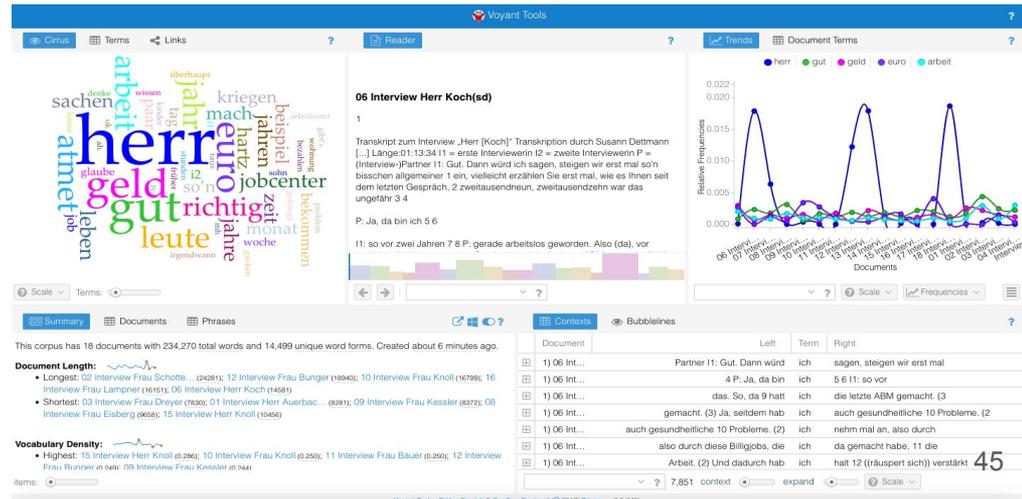
# Sekundäranalyse



- Anlegen eines eigenen Projekts für eine Sekundäranalyse
- Hinzufügen von relevanten Dokumenten aus existierenden Studien
- Kontextualisierung
- Verlinkung von Inhalten mit anderen Inhalten (Text + Bild)
- Kollaboratives Arbeiten
- Möglichkeit zum Teilen (Sharing) mit anderen Wissenschaftlern
- Zusätzliche Interviews
- Möglichkeit der Rückführung von Ergebnissen in die Sammlung
- Kodierung von Inhalten über etablierte Werkzeuge wie MaxQDA, ATLAS.ti

# Resultierende Werkzeuge + Komponenten

- Werkzeug zum Verwalten der eigenen Projekte
- Werkzeuge des TextGrid Labs zur
  - Referenzierung innerhalb der Daten
  - Werkzeug zur Kommentierung / Annotation von Texten
  - Anbindung an bestehende NLP-Werkzeuge, wie z.B. Voyant, zur Exploration der gespeicherten Daten



# Conclusion Sekundäranalyse

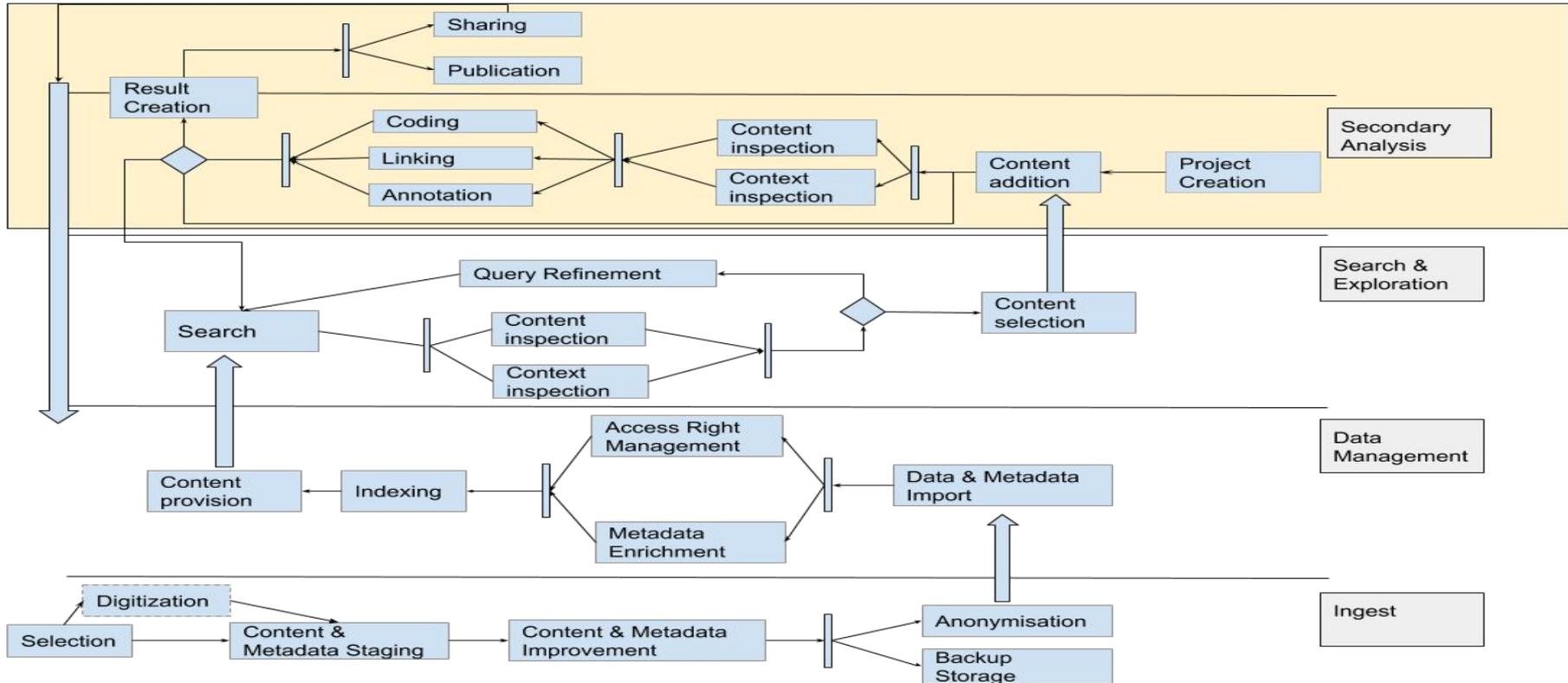
Wo sind die Baustellen?

- Sekundäranalyse als Neuland für alle (bisheriger Schwerpunkt auf Basiskomponenten)

Was ist gelöst?

- Testsysteme, wie z.B. Voyant, aufgesetzt und diese befinden sich zurzeit in der Erprobung

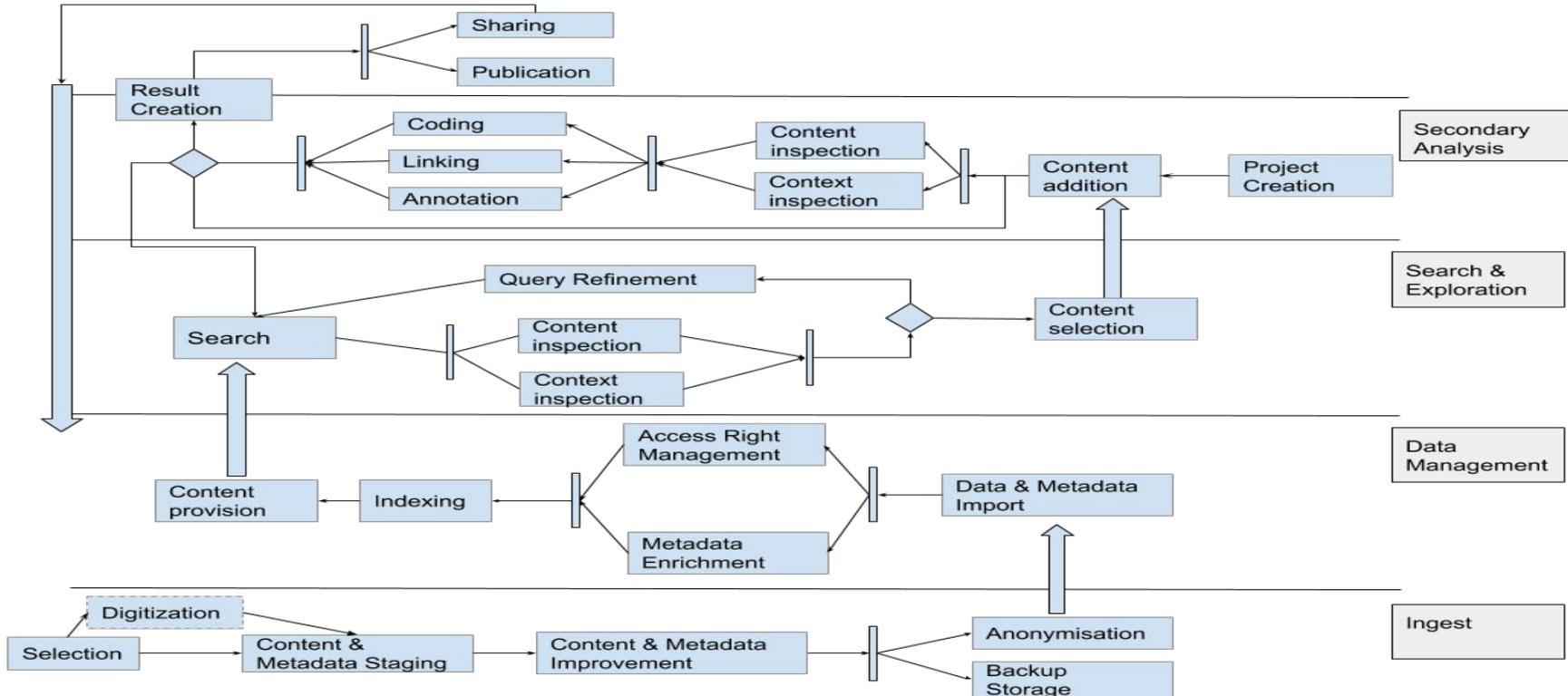
# Kernprozesse: Sekundäranalyse



# Agenda dieser Session

- Einleitung
- Teilprojekt: “Widersprüchliche Integration von Frauen in Dienstleistungsarbeit (WIFE)”
- Prozesse zur Unterstützung der Sekundäranalyse
  - Ingest
  - Datenmanagement und -modellierung
  - Exploration & Suche
  - Sekundäranalyse & Kontextualisierung
- Architektur der eLabour Forschungsinfrastruktur (“Big Picture”)

# Kernprozesse von eLabour

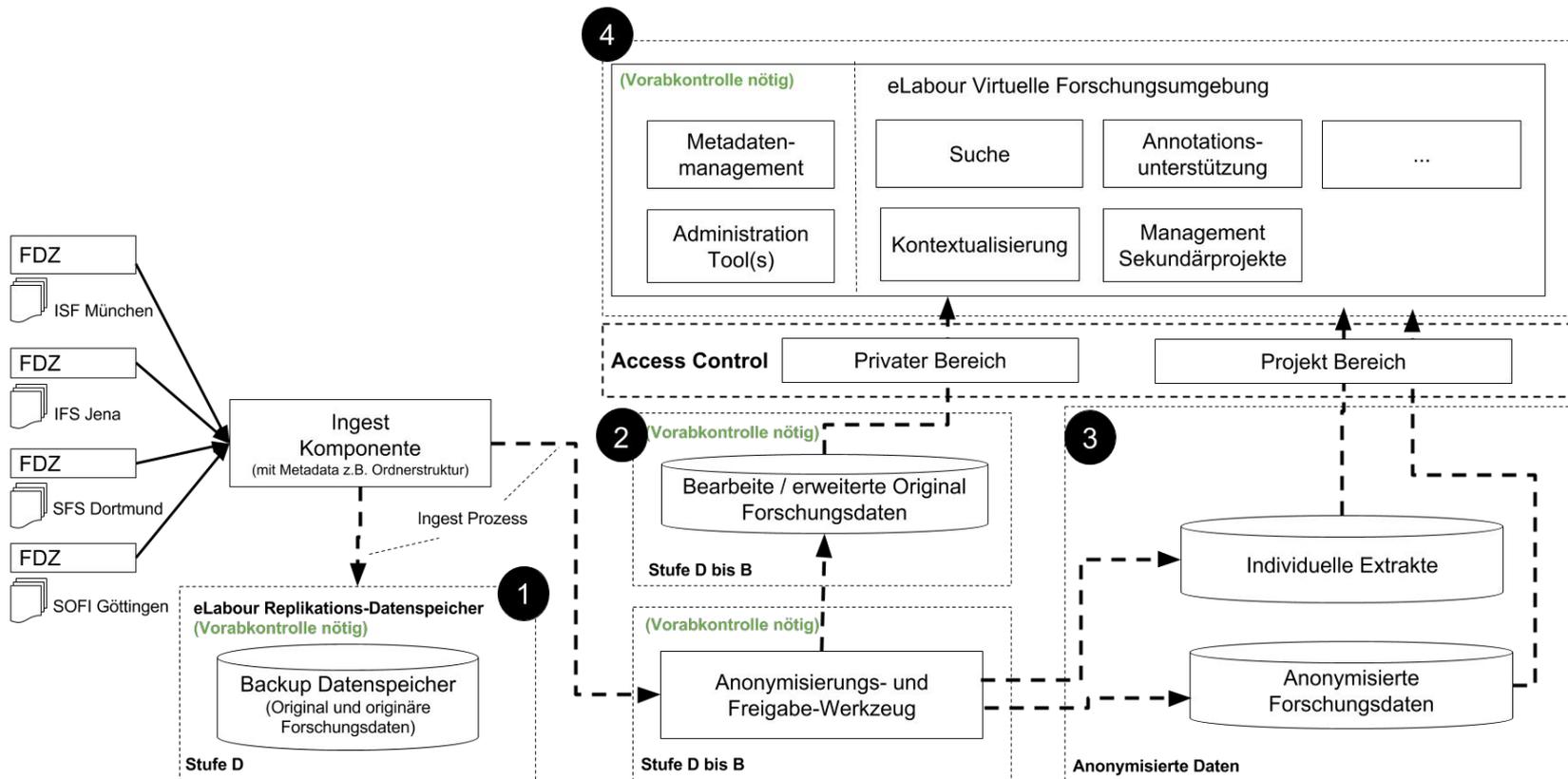


## Umsetzung der Kernprozesse → Forschungsdateninfrastruktur

### Vorgehensweise:

- Erstellung von Anforderungsanalysen, um konkrete Arbeitsschritte und daraus resultierende Arbeitsschritte identifizieren zu können
- Verteilte Entwicklung der Forschungsumgebung durch GWDG, L3S und SUB
- Permanente und dauerhafte Einbindung und Interaktion mit soziologischen Partnern
- Erstellung eines Prototypen und Weiterentwicklung auf Basis der NutzerInnen-Rückmeldungen
- Erstellung eines IT-Konzepts für eine Betriebsphase

# Architektur-Überblick



Herzlichen Dank für Ihre  
Aufmerksamkeit!

Fragen?

*e*Labour