

Claudia Niederée

IT Unterstützung für die Kontextualisierung

Präsentation bei eLabour Workshop
Göttingen, 8. und 9. Februar 2017



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



DLR Projektträger

Kontextualisierung

Arbeitsdefinition:

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

- Erstellungskontext vs. Interpretationskontext
- Schwerpunkt auf zeithistorischer Kontextualisierung



Kontextualisierung als Suchproblem

- Welche Art von Kontext soll betrachtet werden?
- Was soll kontextualisiert werden?
- Was sind relevante Quellen für Kontextinformation?
- Wie müssen diese Quellen vorbereitet werden?
- **Wie findet man relevante Kontextinformation in den Quellen?**
 - **Interdisziplinäre Frage: Was ist relevant als Kontext**
- Wie soll Kontextinformation bereitgestellt werden (z.B. geeignete Granularität)?

Was wird kontextualisiert?

- Primärstudie
 - Was hilft bei der Interpretation der Primärstudie?
 - Unterschiedliche Granularitäten möglich
- Sekundärprojekt
 - Was hilft bei der Interpretation der Primärstudien im Kontext der neuen Fragestellung?
 - Kontextualisierung auf der Grundlagen gewählter Fälle/Studien
 - Kontextualisierung auf der Grundlage einer Beschreibung

- Information über Einstellungen und relevante Ereignisse
 - z.B. European Social Survey Media Claims (soziologisch relevante Ereignisse)
 - Ereignisse aus GDELT oder Wikipedia)
- Statistische Daten (z.B. Beschäftigtenzahlen)
 - z.B. destatis (amtliche Statistiken zu Beschäftigungsformen, Ar
- Information zu Branchen (z.B. Automobilindustrie)
- Tageszeitungen für aktuelle Ereignisse und Entwicklungen (z.B. FAZ)
- Information zu relevanten Gesetzen und Regularien (z.B. Arbeitszeit)
- Arbeitssoziologischen Quellen
 - Gewerkschaftliche Monatshefte (<http://library.fes.de/gmh/jahresin.html>)
- ...



Aktuell: Sammlung von Quellen auf der Grundlage
der Sekundärstudien im Projekt in
AG Kontextualisierung

Beispiel: Gewerkschaftliche Monatshefte

eLabour

- Von 1950 – 2004 digital verfügbar
- Granularität: Einzelne Beiträge als PDF
- Zeithistorische Diskussion zu Gewerkschaften, Beschäftigung, Organisation von Arbeitern, etc.

Gewerkschaftliche
Monatshefte

Lutz Reyher

Zur Entwicklung der Arbeitslosigkeit in den 80er Jahren*

Lutz Reyher, geboren 1927 in Berlin, Diplom-Volkswirt, studierte Volkswirtschaft in Hannover und Berlin. Er ist Leiter des Arbeitsbereichs „Kurzfristige Arbeitsmarktanalyse“ im Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit in Nürnberg.

Aus Ausgabe 10/81

Ende August 1981 zählten die Arbeitsämter in der Bundesrepublik fast 1,3 Mio. Arbeitslose. Das waren 1½ mal soviel wie vor einem Jahr. Man muß diese Statistik fast 30 Jahre zurückverfolgen, um — 1952! — eine nur etwas höhere Zahl von

Bespiel: ESS Media Claims

- European Social Survey (ESS): Studie aus verschiedenen Ländern um sich änderende öffentlichen Einstellungen und Werte in Europa systematisch zu erfassen
- Angereichert mit “Media Claims” – Beschreibung relevanter Ereignisse, die Situation und Einstellungen verändern
- Interessant da bereits mit soziologischen Konzepten assoziiert
- Daten ab 2002

- **Frage:** On the whole how satisfied are you with the present state of the economy (2008) (ESS5 dataset <http://www.europeansocialsurvey.org/>)
- **Germany Officially Sinks Into Recession:** “Germany, Europe's biggest economy and the world's top exporter, has fallen into recession, official figures show Thursday. The German economy contracted by 0.5 percent in the third quarter, according to data released by the national statistics service Desatis on Thursday, November 13. This was more than expected and comes on the heels of a revised retraction of 0.4 percent in the second quarter” (Events concerning the national economy, labour market)

Grad der Automatisierung

Stufe 1: Quellensammlung

- Sammlung und Beschreibung von relevanten Kontextualisierungsquellen als Teil der Infrastruktur
- Kollaborative Erweiterung
- Suche nach Kontextualisierungsquellen

Stufe 2: Anbindung

- Anbindung von Suchunterstützung für ausgewählte Quellen
- Nur für digitale Quellen möglich

Stufe 3: Integration

- Aufbereitung relevanter Quellen für Kontextualisierung (Anreicherung)
- Vorbereitung durch Informationsextraktion, Fragmentierung, Indizierung
- Suchunterstützung

Automatisierung

Stufe 4: Kontextualisierungsvorschläge

- Vorschlag von Kontextanfragen/Kontextfragmenten
 - Extraktion von Information aus zu kontextualisierendem Dokument (Zeit, Themen, Branche)
 - Auf Grundlage des Kontextualisierungsmodells

Noch offen: Wieviel Automatisierung ist möglich?
Wieviel Automatisierung ist gewünscht?

- Technische Vorbereitung
 - Digitale Verfügbarkeit und Formate prüfen
 - Herunterladen oder Verlinken?
 - Unterschiedliche Arten von Quellen berücksichtigen
 - z.B.: Text vs. Daten
 - Lange vs. kurze Texte (z.B. Buch)
 - Quellenspezifisch
- Erschließung
 - Fragmentierung
 - Anreicherung
 - Indizierung

Beispiel: statistischen Daten

- Herunterladen geeigneter Daten von
- Beschreibung der Daten mit relevanter Information für die Kontextualisierung (z.B. Zeit, Art der Daten, Branchenbezug)
- Indizierung der Beschreibungen

Prozess der dynamische Kontextualisierung (Stufe 4)

- Vorbereitung der Kontextquellen
- Anreicherung und Indizierung
- Generierung von Kontextanfragen
- Kontextkandidaten
- Ranking von Kontextkandidaten
- → Vorschläge für Kontextinformation

- Größte Herausforderung der Kontextualisierung
 - Welche Quellen sollen im Einzelfall genutzt werden?
 - Welche Kontextfragmente aus diesen Quellen sind relevant?
- Ansatzpunkte
 - Zeitliche Bindung: Zeitraum der Primärstudie, Zeitraum, den Sekundärprojekt abdecken soll
 - Inhaltlich: Thema der Studie (semantische Suche), Thema des Sekundärprojekts
 - Relevante Branchen
 -
- Erkennen was sich verändert hat
- Noch weitere Verbesserung des interdisziplinäre Kontext nötig

- Weitere Entwicklung des Kontextualisierungsbegriffs in der AG Kontextualisierung; Wichtige Fragen:
 - Welche Arten von Kontextinformation wird benötigt?
 - Welche Kontextquellen sind dafür relevant?
- Sammlung und technische Sichtung der Kontextquellen (getrieben durch Projekte zur Sekundäranalyse)
- Experimente mit ausgewählten Quellen

Herzlichen Dank für Ihre Aufmerksamkeit

Weitere Informationen:

Claudia Niederée
niederee@L3S.de

Nam Tran
ntran@L3S.de