

„Gute Arbeit“ nach dem Boom. (Re-SozIT)

Ein Pilotprojekt zur Längsschnittanalyse arbeitssoziologischer Betriebsfallstudien mit neuen e-Humanities-Werkzeugen

Durchführende Forschungseinrichtungen

Prof. Nicole Mayer-Ahuja
Soziologisches Forschungsinstitut Göttingen
an der Universität Göttingen

Prof. Wolfgang Nejdil
Forschungszentrum L3S
Leibniz Universität Hannover

Prof. Lutz Raphael
Lehrstuhl für Neuere und Neueste Geschichte
Universität Trier

Disziplinäre Verortung

Soziologie, Zeitgeschichte und Informatik

Wissenschaftlicher Use case

Die Arbeitswelt hat sich seit Ende der prosperierenden Nachkriegsjahrzehnte grundlegend verändert – doch wie wirkte sich die neue Konstellation nach dem ökonomischen Boom auf Vorstellungen von „guter Arbeit“ aus? Die Beantwortung dieser Frage erfordert zum einen den Bruch mit der in der Arbeitssoziologie dominierenden Beschränkung auf empirische Momentaufnahmen, weil zu verschiedenen Zeitpunkten erhobenes Material in einer Längsschnittanalyse aufeinander bezogen werden muss, um Veränderungen im Zeitverlauf analysieren zu können. Zum anderen ist die methodologische Herausforderung zu meistern, wie die im Rahmen qualitativer Sozialforschung generierten Forschungsmaterialien aus mehreren Jahrzehnten einer Sekundäranalyse mit e-Humanities Werkzeugen zugänglich gemacht werden können. Angestrebt wird dabei eine (bislang kaum praktizierte) soziologische und geschichtswissenschaftliche Kontextualisierung arbeitssoziologischer Fallstudien.

Für das Pilotvorhaben steht das umfangreiche Forschungsdatenarchiv des SOFI mit etwa 50 arbeits- und industriesoziologischen Projekten seit den 1970er Jahren zur Verfügung. Um dieses Material für Primärprojekte übergreifende Sekundäranalysen zu erschließen, werden IT-Werkzeuge entwickelt und erprobt, die es erlauben, systematisch und kollaborativ, entwicklungsorientiert und reflektierend auf dieses Material zuzugreifen und Analysen des qualitativen empirischen Materials durchzuführen. Für diese - teils semi-automatische – Analyse werden Ansätze wie Topic Modelling, Sentiment-Analyse, etc. aus den Bereichen Information Retrieval und Text Mining aufgegriffen und angepasst.

Methode

In enger Kooperation zwischen den soziologischen und zeithistorischen und Informatik-Teilprojekten wird das qualitative Datenmaterial der SOFI Studien aus vier Jahrzehnten erschlossen und der soziologischen und zeitgeschichtlichen Sekundäranalyse zugänglich gemacht. Die in den arbeitssoziologischen Primärstudien von 1966 bis heute erhobenen empirischen Momentaufnahmen sollen systematisch aufeinander bezogen werden um Veränderungen im Zeitverlauf analysieren zu können. Ausgehend von der sekundäranalytischen Fragestellung werden mit neuen IT-Werkzeugen Interviews gesucht und ausgewertet, die hierzu Aussagen enthalten. Allerdings können diese Interviews nicht losgelöst von ihrem historischen und empirischen Entstehungskontext analysiert werden, vielmehr müssen sie im jeweiligen Kontext der Primärstudien erschlossen und analysiert, d.h. (re-)kontextualisiert werden. Bei dieser soziologischen und geschichtswissenschaftlichen Re-Kontextualisierung qualitativer, arbeitssoziologischer Studien werden methodisch neue Wege

beschritten, um die im Rahmen qualitativer Sozialforschung in vier Jahrzehnten generierten Materialien aus einer Vielzahl von Primärstudien für Längsschnittanalysen zu erschliessen.

Sekundäranalysen basierend auf Studien aus dem SOFI seit den 1960er Jahren (Längsschnitt)

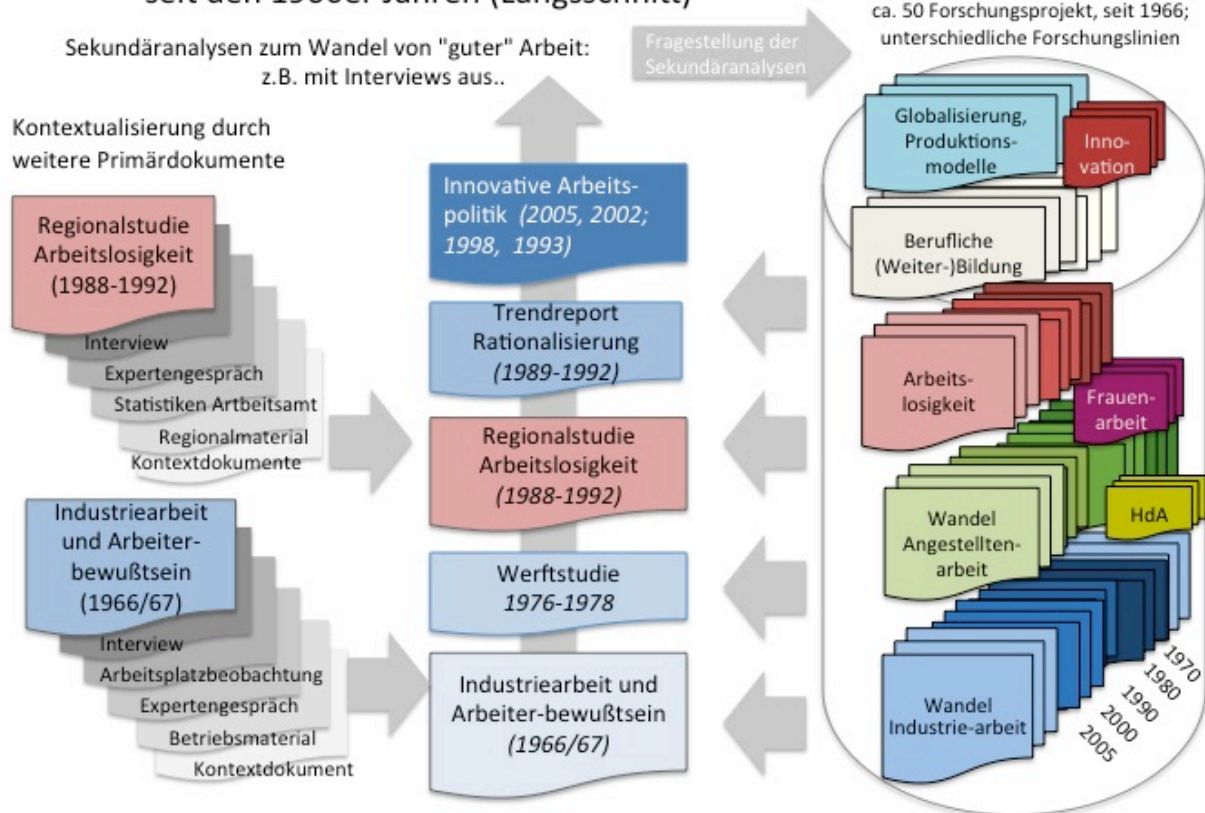


Abb. Projektübergreifende Sekundäranalysen im Längsschnitt

Hierzu werden IT-Werkzeuge entwickelt und eingesetzt, die es erlauben, systematisch, kollaborativ und reflektierend primärstudienübergreifende Längsschnittanalysen mit dem umfangreichen empirischen Material aus dem digitalen Forschungsarchiv des SOFI durchzuführen.

Die IT-Werkzeuge lassen sich in drei Kategorien unterteilen:

- 1. Intelligente Exploration textueller Daten.** Für den effektiven Zugriff auf umfangreiche, heterogene Forschungsdaten ist eine intelligente Volltextsuche erforderlich, die durch Instrumente zur Navigation (Facettensuche) und weitere Maßnahmen zur Anreicherung und Strukturierung von Suchergebnissen und Visualisierungen wie beispielsweise Zeitlinien, Wortwolken, etc. ergänzt wird.
- 2. Extrahieren neuen Wissens und Aggregieren von Informationen.** Durch Einsatz und Weiterentwicklung von Verfahren zu Informationsextraktion und Text Mining werden zusätzliche Informationen gewonnen, bspw. in Form von (semi-)automatisch erzeugten Annotationen zu extrahierten Entitäten, Themen (Topic Modelling) und Einstellungen (Sentiment-Analyse).
- 3. Kontextualisierung von Inhalten.** Die Sekundäranalyse von Studien und einzelnen Interviewpassagen setzt den Zugriff auf zusätzliche Erhebungs- und Kontextinformation voraus. Die Längsschnittanalysen erfordern, nicht nur einzelne Textpassagen in ihrem historischen Kontext zu verorten, sondern auch Primärdaten aus verschiedenen Jahren zueinander in Verbindung zu setzen und dabei den sich verändernden Kontext jeweils zeitspezifischer Vorstellungen von „guter“ Arbeit mit zu reflektieren. Basierend auf Text Mining-Strategien und ggf. unter Zuhilfenahme externer Wissensquellen soll eine Kontextanreicherung erreicht werden.

5000 mal 5000: VW startet in eine neue Arbeitswelt ...
 Zu den 3500 Beschäftigten, die bis Mai eingestellt werden, kommen noch etwa 150 Meister und Facharbeiter dazu. In der Sprache der 5000er Fabrik sind das Betriebsingenieure und sogenannte Fachtalente. Sie unterstützen die angeleiteten Talente und helfen in kniffligen Situationen aus.

Kontextualisierung

Anonymisierung

Opinion Mining /
Sentiment-Analyse

Informationsextraktion
z.B. Topic Modelling

Sofi: Ist das unter dem Gesichtspunkt Gerechtigkeit richtig?
 Herr ■ Ja, zwischen Talenten und Fachtalente würde ich da keinen Unterschied machen, wobei ich mir schon ein bisschen Gedanken mache, wenn ich an die Qualifizierungsmaßnahmen der Fachtalente denke, die müssen meiner Ansicht nach noch ein bisschen mehr leisten und da hätte ich es auch akzeptiert, wenn die ein bisschen mehr kriegen. Ich habe eben nicht die Ausbildung, die die gemacht haben. Dass die Betriebsingenieure mehr kriegen, ist klar.
 Sofi: Es gibt ja bei ■ keine Meister im klassischen Sinn mehr, jetzt hat man die Betriebsingenieure, wie beurteilen Sie das?
 Herr ■ Ich finde es gut. Es muss immer Leute geben, die über einem stehen, wenn es zu Konflikten oder Problemen kommt, die im Team nicht mehr lösbar sind. Man erwartet irgendwo auch eine gewisse Führung, man kann die Leute nicht einfach machen lassen was sie wollen. Ich denke, das ist in Ordnung mit den Betriebsingenieuren. Ich hoffe nur, dass die auch entsprechend menschlich rüberkommen und nicht nur auf ihre Chefposition pochen. Aber eigentlich gehe ich nicht davon aus, ich denke, dass die auch von der Mentalität her ähnlich sein werden und ähnlich rüberkommen, dass aber jedem trotzdem bewusst ist, dass er der Vorgesetzte ist. Ich glaube nicht, dass der darauf bestehen muss.

Intelligente Suche & Exploration



Abb. IT-Werkzeuge für die Sekundäranalyse von Interviews aus qualitativen soziologischen Studien

In kurzen Entwicklungszyklen werden die beschriebenen IT-Verfahren implementiert, auf ihre Eignung hin evaluiert, weiterentwickelt und praktisch zur Beantwortung der soziologischen und historischen Fragestellungen rund um „gute“ Arbeit eingesetzt. Die Nützlichkeit der Werkzeuge kann so bereits während der Entwicklung in der Anwendung durch die SoziologInnen und HistorikerInnen erprobt und verbessert werden. Die WissenschaftlerInnen aus Arbeitssoziologie und Zeitgeschichte wiederum können schnell auf IT-Werkzeuge zugreifen, die speziell auf ihre Anforderungen hin entwickelt werden.

Im Hinblick auf die Problematik der Datenanonymisierung wird zudem ein Konzept erstellt, welches auf den Bausteinen Risikoklassifizierung, IT-gestützte Anonymisierungsmaßnahmen und infrastrukturelle Rahmenbedingungen mit abgestuften Zugangsrechten beruht.

Genutzte Ressourcen, existierende Forschungsdaten

Daten:

Forschungsdaten aus dem Archiv des Soziologischen Forschungsinstituts

- Dokumente aus ca. 50 Forschungsprojekten, die am SOFI seit Ende der 1960er Jahre durchgeführt und archiviert wurden;
- davon ca. 250.000 nachträglich digitalisierte Dokumente aus Projekten ab 1968 bis Mitte der 1990er Jahre; überwiegend Leitfadeninterviews mit Beschäftigten, Expertengespräche, Dokumentationen von Zwischenergebnisse und unterschiedlichen Kontextinformationen;
- digital erstellte Dokumente aus Projekten seit Mitte der 1990er Jahre

Verfahren:

WissenschaftlerInnen aus Arbeitssoziologie und Zeitgeschichte verarbeiten die Forschungsdaten im SOFI Archiv mit Hilfe der Werkzeuge, die von der Informatik entwickelt werden:

- Auswahl der für die Fragestellung der Längsschnittanalysen benötigten Projekte durch die beteiligten SoziologInnen und HistorikerInnen (unterstützt durch Mitarbeiter des SOFI)
- IT-gestützte Erschließung der für die Sekundäranalysen verwendbaren Dokumente aus unterschiedlichen Projekten
- Erstellen einer geeigneten Datenstruktur, Aufbau eines Pilotarchivs für die Entwicklung der IT-Werkzeuge (unter Berücksichtigung von Metadatenstandards). Im Backend: Fedora und XML. Basisfunktionen zur Annotation, Verlinkung, Import, Export.

- Entwicklung, Test und Anwendung der Werkzeuge zur Exploration, Extraktion und Aggregation der Daten für die Längsschnittanalysen. Dabei werden Open Source Tools wie bspw. Mallet, UIMA, Gate oder LingPipe, WEKA eingesetzt und adaptiert werden.
- Entwicklung und Anwendung von Methoden und Tools zur (Re-)Kontextualisierung von (extrahierten) Forschungsdaten.
- Entwicklung eines IT-gestützten Anonymisierungsansatzes für qualitative arbeitssoziologische Forschungsdaten unter Berücksichtigung von aktuellen Ansätzen bspw. zur Datenaggregation mit k-Anonymity.

Entstehende Ressourcen

- Pilotdatenbank mit qualitativen Forschungsdaten aus ausgewählten Projekten des SOFI Archivs; sie wird nach Projektablauf für die Nutzung durch andere WissenschaftlerInnen zur Verfügung gestellt werden (unter Berücksichtigung der Anonymisierungsregeln)
- die zu entwickelnde (Meta-)Datenstruktur und das Datenbankmodell sind Grundlage für den weiteren Aufbau des digitalen SOFI-Archivs, das unter Berücksichtigung der besonderen Datenschutzerfordernungen ebenfalls für wissenschaftliche Nutzung zur Verfügung stehen soll (unter Berücksichtigung der Vorschläge für Non-Disclosure-Agreements)
- IT-Werkzeuge für die Erschließung und (semi-)automatische Analyse heterogener qualitativer Forschungsdaten, insbesondere für Exploration, Extraktion und Aggregation von Daten, sowie für deren Kontextualisierung. Ausgewählte Werkzeuge werden als Open Source zur Verfügung gestellt.
- Guidelines / Best Practices zur Anonymisierung qualitativer Forschungsdaten am Beispiel des SOFI-Archivs

Kooperationen

- Göttinger Centre für Digital Humanities
- DARIAH
- Datenservicezentrum Betriebs- und Organisationsdaten (DSZ-BO), Universität Bielefeld
- Archiv für Lebenslaufforschung (ALLF) / eScience lab, Universität Bremen
- angestrebt sind Kooperationen hinsichtlich Metadatenstandards für qualitative Forschungsdaten (zusätzlich zu den o.g., SFB 600 Universität Trier, DDI, u.a.)
- angestrebt ist eine Kooperation mit ATLAS.ti über ein Datenformat für Import und Export semi-automatisch annotierter Forschungsdaten
- Kooperation mit weiteren Anwendern, z.B. dem Institut für Sozialforschung (ISF) München und dem International Institute for Social History (IISG) in Amsterdam

Kontaktinformationen

Projektwebseiten:

SOFI <http://www.sofi-goettingen.de/index.php?id=1086>

L3S <http://www.l3s.de/GuteArbeit>

Projektleitung:

Prof. Nicole Mayer-Ahuja
Universität Hamburg / SOFI Göttingen
Nicole.Mayer-Ahuja@wiso.uni-hamburg.de

Kontakt Team Neuere Zeitgeschichte:

Dr. Jenny Pleinen
Lehrstuhl Raphael
pleinenj@uni-trier.de

Projektkoordination und Kontakt SOFI:

Heidmarie Hanekop
SOFI Göttingen
heidi.hanekop@sofi.uni-goettingen.de

Kontakt L3S Team:

Dr. Claudia Niederée
Forschungszentrum L3S, Hannover
niederee@l3s.de