

Holleriths Vermächtnis – ein Beitrag zur Geschichte von Frauen in der EDV

Topic Modeling als Methode digitaler Sekundäranalyse lebensgeschichtlicher Interviews

Dennis Möbus

Lebensgeschichte und Sozialkultur im Ruhrgebiet (LUSIR) war das erste große Oral-History-Projekt in Deutschland. 1980 von Lutz Niethammer initiiert, entstand eine Sammlung von etwa 350 biographischen Interviews, die von einer Gruppe von Forscher*innen ausgewertet wurden (Niethammer 1983a; Niethammer 1983b; Niethammer/von Plato 1985). Heute sind die Interviews im Archiv „Deutsches Gedächtnis“ des Instituts für Geschichte und Biographie der FernUniversität in Hagen verfügbar und werden derzeit für ein Onlinearchiv digital aufbereitet.

Kern der Interviewsammlung sind die Lebensgeschichten von Arbeiter*innen und Angestellten, die zwischen 1930 und 1980 im hochindustrialisierten Ruhrgebiet gelebt haben. War das Erkenntnisinteresse ursprünglich die Erfahrung bzw. Verarbeitung von Faschismus und Weltkrieg, zeichnen die umfangreichen Stegreiferzählungen weit darüber hinaus ein Bild der Lebenswelten im westdeutschen Industrierevier der Nachkriegszeit. Es kommen nicht nur Berg- und Stahlarbeiter zu Wort, auch Frauen in Verwaltungsberufen, leitende Angestellte und Gewerkschaftsfunktionäre sind in der LUSIR-Sammlung vertreten. Durch den biographischen Ansatz der Oral History bietet jede Lebenserzählung eine ganz eigene Innensicht auf Jahrzehnte sozialen Wandels zwischen Weimarer Republik, Nationalsozialismus und Bundesrepublik. So finden sich neben den für die Primärforschung ursprünglich zentralen „Faschismuserfahrungen“ auch Wahrnehmungen des sogenannten „Wirtschaftswunders“ und des „Zeichensterbens“, was Einblicke in tiefgreifende und zunächst widersprüchlich erscheinende Modernisierungs- und Deindustrialisierungsprozesse gewährt. Diese Einblicke konnten durch eine computergestützte Sekundäranalyse der LUSIR-Interviews teils gewonnen, teils vertieft werden – immer der Philosophie der Oral History folgend, nicht nur historische Fragen zu beantworten, sondern auch zur Formulierung neuer Fragen beizutragen. Mit Hilfe von Topic Modeling, einem Text-Mining-Verfahren zur Erschließung großer unstrukturierter Datenbestände, konnten in den digitalisierten Transkripten Erzählungen über Automatisierungs- und Computerisierungsprozesse ausfindig gemacht werden. Von Lochkarten- oder Hollerithmaschinen ist da die Rede oder auch von frühen Computern und der CNC-Technologie.

In diesem Aufsatz sollen zwei Perspektiven auf die Digitalisierung eingenommen werden: Auf inhaltlicher Ebene wird die Wahrnehmung früher Digitalisierungsprozesse von vor allem weiblichen Industriearbeiter*innen und -angestellten in den Blick

genommen. Auf methodologischer Ebene soll beispielhaft geprüft werden, wie die digitale Repräsentation von Oral-History-Interviews und deren computergestützte Analyse zur Geschichtsschreibung sozialen Wandels in Westdeutschland beitragen können. Der Beitrag versteht sich als explorative Studie, die das Potential eines solchen methodischen Ansatzes ausloten will.

Oral History digital

Das Archiv „Deutsches Gedächtnis“ umfasst gut 3.000 Interviews aus über einhundert Forschungsprojekten, darunter die rund 350 Interviews aus dem LUSIR-Projekt. Um diese wertvollen Lebensgeschichten und wichtigen historischen Quellen langfristig zu erhalten und komfortabel zugänglich zu machen, arbeitet das Archiv mit Forschungsinstituten aus der Informatik zusammen.

In einem gerade abgeschlossenen Projekt konnten in Zusammenarbeit mit dem *Fraunhofer-Institut für intelligente Analyse- und Informationssysteme (IAIS)* entscheidende Fortschritte in der automatischen Spracherkennung (ASR) von Interviews gemacht werden, was wegen der mangelhaften Aufnahmequalität und spontaner, dialektaler oder undeutlicher Aussprache eine große Herausforderung war (Leh et al. 2018).¹ Die Audiomining-Plattform des Fraunhofer-Instituts kann von Mitarbeiter*innen und Nutzer*innen des Archivs „Deutsches Gedächtnis“ genutzt werden, um Transkripte der Interviews automatisiert zu erstellen. Darüber hinaus erlaubt das Audiomining über das Time Alignment Volltextsuchen auf kompletten Sammlungen mit direktem Zugang zum Audiosignal und ermöglicht die Auswertung einer großen Zahl von Datensätzen.

Im jüngsten Projekt, „Oral-History.Digital“, das vom *Center für Digitale Systeme (CeDiS)* an der Universitätsbibliothek der Freien Universität Berlin geleitet wird, entsteht eine webbasierte Informationsinfrastruktur für audiovisuelle lebensgeschichtliche Interviews. Diese Umgebung wird die Onlinerecherche in Oral-History-Sammlungen verschiedener Archive sowie das Indexieren und Annotieren der Interviews ermöglichen.² Eine Weiterentwicklung, die auf diesen beiden Säulen aufsetzt, wäre eine computergestützte inhaltliche Analyse der digitalisierten, aber unstrukturierten Daten, um diese systematisch zu erschließen und eine bessere Durchsuchbarkeit zu gewährleisten.

Vom Zählen zum Schätzen: Minenarbeit in historischen Dokumenten

Als Beispiel einer solchen Analyse sollen hier die Ergebnisse einer Auswertung der LUSIR-Interviews mit Hilfe von basalen Text-Mining-Algorithmen und des Topic Modeling präsentiert werden. In der Audiomining-Plattform des Fraunhofer-Instituts ist bereits eine Keyword-Extraktion implementiert, die besonders häufige Wörter auflistet und einen ersten Eindruck der reichhaltigen Inhalte lebensgeschichtlicher Interviews vermittelt. Um diese Stichwortliste zu verfeinern, wurde ein einfacher Auszählalgorithmus in der Programmiersprache Python geschrieben. Mit Hilfe einer selbst ge-

1 Das Projekt „KA³ - Kölner Zentrum Analyse und Archivierung von AV-Daten“ unter Leitung von Nikolaus Himmelmann wurde von 2015 bis 2020 durchgeführt und vom BMBF gefördert: <https://www.forschungsdaten.info/fdm-im-deutschsprachigen-raum/nordrhein-westfalen/disziplinspezifische-angebote/ka3/> (18.11.2021).

2 Das Projekt wird von der DFG gefördert. Näheres unter: <https://www.oral-history.digital/> (25.11.2021).

pflegten Stoppwortliste, die neben Artikeln und Pronomen auch extrem häufige Begriffe enthält, konnten die Ergebnisse gefiltert und nochmals deutlich verbessert werden. Dazu wurden beim automatischen Durchzählen der als unformatierte Textdateien vorliegenden Transkripte die in dieser Stoppwortliste enthaltenen Wörter ignoriert und die übrigen, vermeintlich sinntragenden Begriffe der Häufigkeit nach aufgelistet. Es ist wenig verwunderlich, dass sich die häufigsten Keywords um Kriegserfahrungen und die industrielle Arbeitswelt drehen: „Zeit“, „Jahre“, „Krieg“, „Arbeit“ und „Kinder“. Auch der Blick auf weniger häufige Begriffe hält keine Überraschungen bereit, doch immerhin geben sie die hauptsächlichen Inhalte der Sammlung recht gut wieder: „Krupp“, „Gewerkschaft“, „Betriebsrat“, „Zeche“ oder „SPD“ tauchen unter den 100 häufigsten Begriffen auf. Ein Hinweis darauf, dass die automatisierte Generierung und Gewichtung von Keywords für die Erschließung unbekannter Sammlungen ein durchaus taugliches Instrument darstellen.

Eine simple, aber hilfreiche Methode, um den Inhalt unstrukturierter Texte zu umreißen, sind *n*-Gramme. Über Webservices (z.B. Voyant Tools), freie Software (z.B. KNIME) oder mit wenigen Zeilen Code kann man Wortpaare (Bigramme) oder -tripel (Trigramme) eines Textes nach Häufigkeit – bereinigt von Stoppwörtern – ausgeben.³ Unter diesen Bi- und Trigrammen erscheinen Namen von Managern, Industriellen oder Institutionen mit Bezug zum Ruhrgebiet, unter letzteren beispielsweise das Unternehmen „Bochumer Verein“ oder die Gewerkschaft „IG Metall“. Neben diesen *Named Entities* finden sich etwas weiter hinten in den Top-100 auch historisch signifikante Aussagen wie „Krieg verloren“ (36 Erwähnungen). Betrachtet man die Trigramme, erregt die Phrase „deutschen Sieg geglaubt“ (19 Erwähnungen) Aufmerksamkeit.

In einem nächsten Schritt wurde nach Geschlechtern differenziert. Es zeigt sich, dass Frauen – entgegen der Darstellungen der Nachkriegsjahre im LUSIR-Projekt selbst (vgl. Schmidt 1983b: 191) – durchaus über Politik und Arbeit sprechen. Entsprechende Keywords und Phrasen sind jedoch weniger häufig als bei Männern, dafür verweisen Begriffe wie „Kinder/Kind“, „zu Hause“, „Tochter“, „Sohn“ und „geboren“ auf die Dominanz häuslich-familiärer Themen (vgl. Schmidt ebd.: 200).

Um einen klareren Eindruck von den Narrativen zu gewinnen, ist die Anwendung anspruchsvollerer Text-Mining-Verfahren erforderlich. Als Methode, die auch auf verhältnismäßig kleinen Datensätzen arbeitet, ist das Topic Modeling eine gute Wahl.⁴ Topic Modeling ist ein probabilistisches Verfahren zur automatisierten Inhaltsanalyse, das auf bayesscher Statistik aufsetzt und aus einem Textabschnitt inhaltsdefinierende Keywords extrahiert. Datengrundlage ist ein Korpus von Dokumenten, das in einem *Preprocessing* digital aufbereitet wird. In unserem Fall wurden etwa 150 digital vorliegenden Interviewtranskripte in kleinere Abschnitte unterteilt, um einen differenzierten Überblick zu erhalten, der auch marginale Themen einfängt. Für eine Evaluation verschiedener Modelle wurden die Transkripte in Abschnitte zu 500 und zu 1.000 Wörtern zerlegt, die Gesamtmenge aller Einzelabschnitte aller Interviews stellt das Korpus – also das Sample an Dokumenten – dar, welches das Programm verarbeitet. Als nächster Schritt wurde ein *Dictionary* erstellt, in dem jedes Wort, das mindestens einmal im

3 Tutorials zum Schreiben basaler Text-Mining-Routinen finden sich im Web massenhaft, etwa hier speziell für Historiker: <https://programminghistorian.org/en/lessons/counting-frequencies> (20.8.2021).

4 Gute Einführungen in das Topic Modeling bieten das entsprechende Kapitel in Graham et. al 2016, das Online-Tutorial auf <https://programminghistorian.org/en/lessons/topic-modeling-and-mallet> (20.8.2021) und der Aufsatz von Fechner/Weiß 2017.

(5506, 'zeit'),	(243, ('fünfziger', 'jahre')),	(70, ('mark', 'im', 'monat')),
(4373, 'jahre'),	(241, ('bochumer', 'verein')),	(58, ('nsdap', 'dnvp', 'frauenverbände')),
(4006, 'krieg'),	(223, ('halbes', 'jahr')),	(45, ('kraft', 'durch', 'freude')),
(3928, 'arbeit'),	(197, ('50er', 'jahren')),	(41, ('laufe', 'der', 'zeit')),
(3321, 'kinder'),	(165, ('herr', 'dewald')),	(39, ('zeit', 'des', 'nationalsozialismus')),
(3283, 'hause'),	(151, ('herr', 'wolf')),	(37, ('art', 'und', 'weise')),
(3024, 'frauen'),	(151, ('herr', 'stasch')),	(35, ('morgens', 'bis', 'abends')),
(3001, 'herr'),	(146, ('50er', 'jahre')),	(34, ('männer', 'und', 'frauen')),
(2891, 'schule'),	(133, ('kurze', 'zeit')),	(31, ('herr', 'dr', 'hartmann')),
(2854, 'früher'),	(129, ('herr', 'muhs')),	(30, ('abends', 'nach', 'hause')),
(2659, 'eltern'),	(119, ('jahre', 'älter')),	(28, ('laufe', 'der', 'jahre')),
(2600, 'essen'),	(119, ('herr', 'spielhoff')),	(26, ('arbeit', 'und', 'brot')),
(2427, 'sag'),	(111, ('heil', 'hitler')),	(23, ('quietschen', 'im', 'hintergrund')),
(2307, 'jahren'),	(109, ('firma', 'krupp')),	(23, ('gründung', 'der', 'bundesrepublik')),
(2305, 'familie'),	(107, ('herr', 'sprockhoff')),	(22, ('muß', 'ich', 'ehrlich')),
(2146, 'wohnung'),	(107, ('herr', 'küppler')),	(20, ('grund', 'und', 'boden')),
(2101, 'wissen'),	(103, ('fünfziger', 'jahren')),	(20, ('gang', 'und', 'gäbe')),
(1988, 'jahr'),	(93, ('ig', 'metall')),	(20, ('essen', 'und', 'trinken')),
(1934, 'erzählt'),	(92, ('schöne', 'zeit')),	(20, ('brach', 'der', 'krieg')),
(1912, 'haus')	(91, ('herr', 'seeling'))	(19, ('uhr', 'zu', 'hause'))

Abb. 1: Top-20 Keywords, Bi- und Trigramme absteigend sortiert nach absolutem Vorkommen im Text (Quelle: eigene Darstellung).

Gesamtsample vorkommt, eine ID in Form eines zweidimensionalen Vektors erhält – visuell betrachtet wird das Wort sozusagen als Punkt in einem Koordinatenfeld verortet.

Topic Modeling beruht auf der Annahme, dass jedes Thema durch bestimmte wiederkehrende charakteristische Begriffe geprägt wird. Dabei kann für die Berechnung der Topics die Reihenfolge der Wörter in einem Textabschnitt vernachlässigt werden, um den Inhalt zu erschließen – man spricht vom *Bag-of-words*-Modell. Entscheidend ist die Anzahl ebensolcher definierender Begriffe, die gemeinsam in einem Textabschnitt vorkommen. Daher wird für jeden Interviewabschnitt ein *Bag-of-words* erstellt und die enthaltenen Wörter nach ihrer Häufigkeit sortiert.⁵ Die Berechnung der Verbindung zwischen gemeinsam häufig vorkommenden Wörtern und deren Nähe oder Distanz zu Worthäufigkeiten in anderen Textabschnitten erfolgt anschließend auf Grundlage der *Latent Dirichlet Allocation* (LDA), einer probabilistischen Funktion. Für diese Untersuchung wurde das Python-Modul *Gensim* verwendet.⁶ In diesem Verfahren wird in mehreren Durchläufen die Wahrscheinlichkeit inhaltlicher Zusammenhänge durch wiederkehrende Worthäufungen berechnet. Dabei werden die letzten Schätzungen immer wieder mit neuen Schätzungen ins Verhältnis gesetzt, um das Modell nach und nach zu verbessern – ein Vorgehen, das stark an den hermeneutischen Zirkel erinnert.

Als Richtlinie für die inhaltliche Ausdifferenzierung der Ergebnisse kann als weiterer Parameter neben der Größe der Textabschnitte die gewünschte Anzahl der Topics vorgegeben werden. Die optimale Anzahl hängt von Umfang und Diversität des Korpus ab. Nach der erfolgreichen Berechnung eines Modells wird ein einzelnes Topic durch die Wörter repräsentiert, die als am Wahrscheinlichsten für die damit verknüpften Textabschnitte gelten. Ein Topic ist somit eine Abbildung eines Themas oder Diskurses, jedoch keine vollständige Entsprechung.⁷ Im hier geschilderten Versuch wurden Modelle mit zehn, zwanzig und fünfzig Topics berechnet und verglichen (Fechner/Weiß 2017; Blei 2012). Bei der Evaluation der Ergebnisse mit Hilfe des Visualisierungstools *pyLDavis* zeigte sich insbesondere beim Modell mit 50 Topics und Textabschnitten zu je 1.000 Wörtern eine interessante Spur (vgl. Abb. 2).

Das Tool *pyLDavis* stellt Topics als Kreise in einem Koordinatenfeld dar. Der Umfang des Kreises entspricht dem Vorkommen der in diesem Topic gesammelten und als zusammengehörig befundenen Wörter. Je besser die Kreise voneinander getrennt sind, desto besser – weil trennschärfer – sollten die Topics modelliert sein. In Abbildung 2 sieht man die in unserem Korpus gefundenen Topics und deren Verknüpfungen. Die Reihenfolge der *Most Salient Terms* auf der rechten Seite der Abbildung gibt die nach dem LDA-Verfahren herausgefilterten häufigsten und aussagekräftigsten Keywords des Korpus wieder.

5 Vgl. <https://programminghistorian.org/en/lessons/topic-modeling-and-mallet> (13.8.2021).

6 <https://radimrehurek.com/gensim/> (20.8.2021).

7 Vgl. dazu die Diskussionen: <https://tedunderwood.com/2012/04/07/topic-modeling-made-just-simple-enough/>; <https://fortext.net/routinen/methoden/topic-modeling> (13.8.2021).

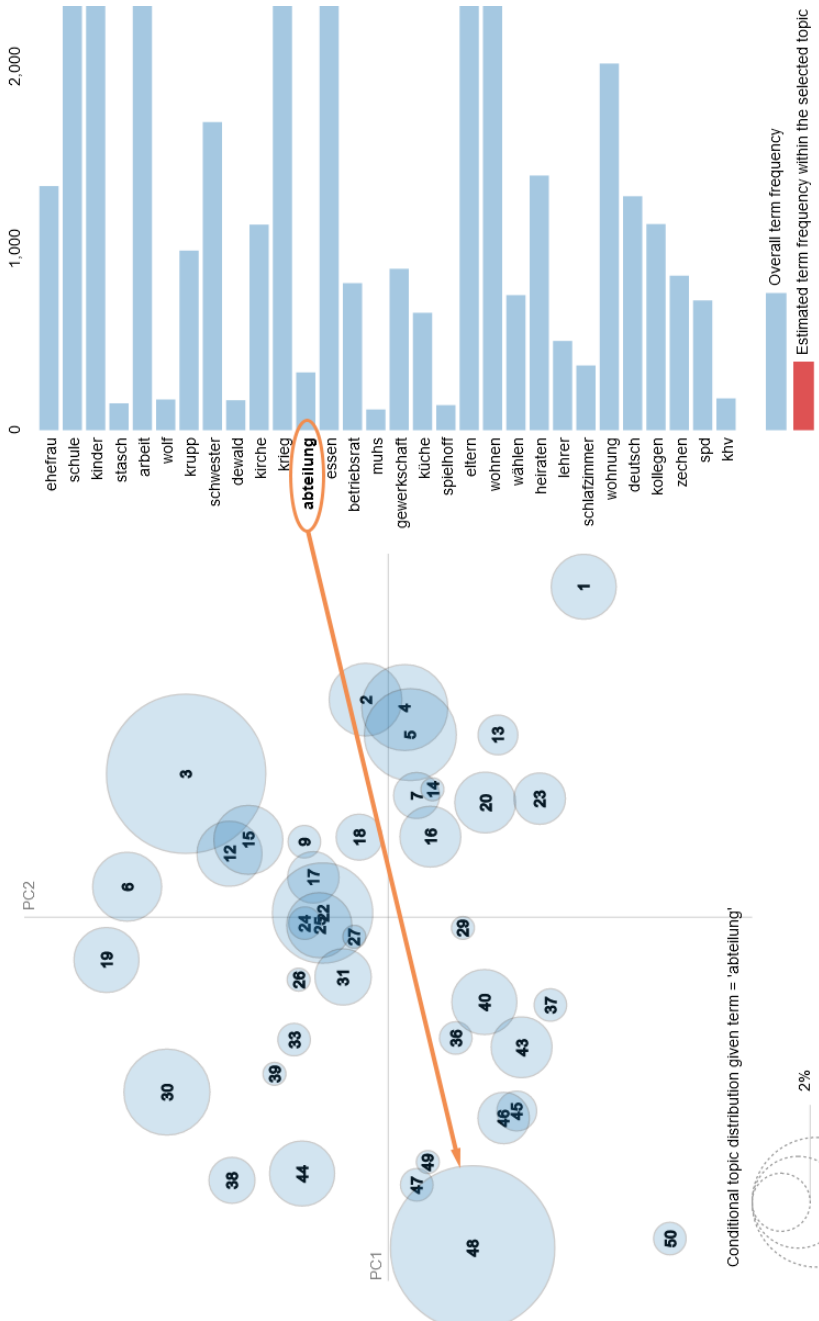


Abb. 2: Übersicht aller 50 Topics als Kreise und der Top-30-Keyworts für das Korpus. Das hervorgehobene Topic 48 ist mit dem Keyword „Abteilung“ verknüpft (Quelle: eigene Darstellung).

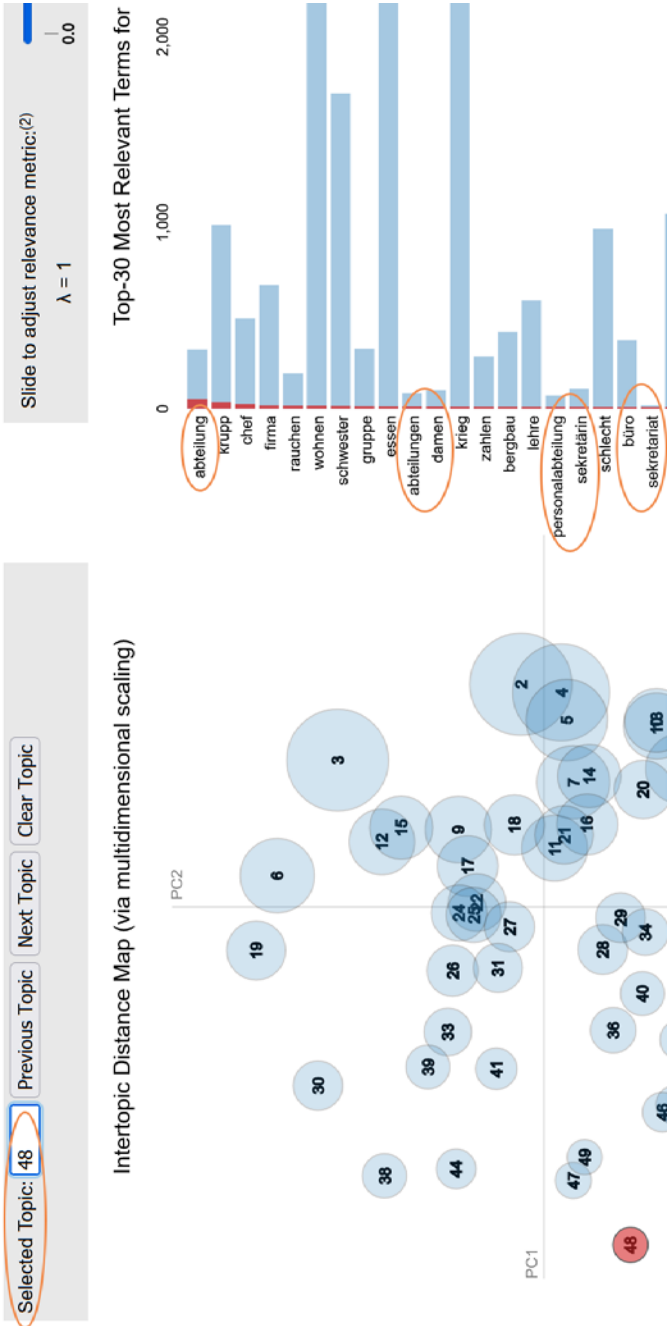


Abb. 3: Auswahl des Topics 48 auf der linken Seite und die laut Topic Modeling aussagekräftigsten Keywords für Topic 48 auf der rechten Seite (Quelle: eigene Darstellung).

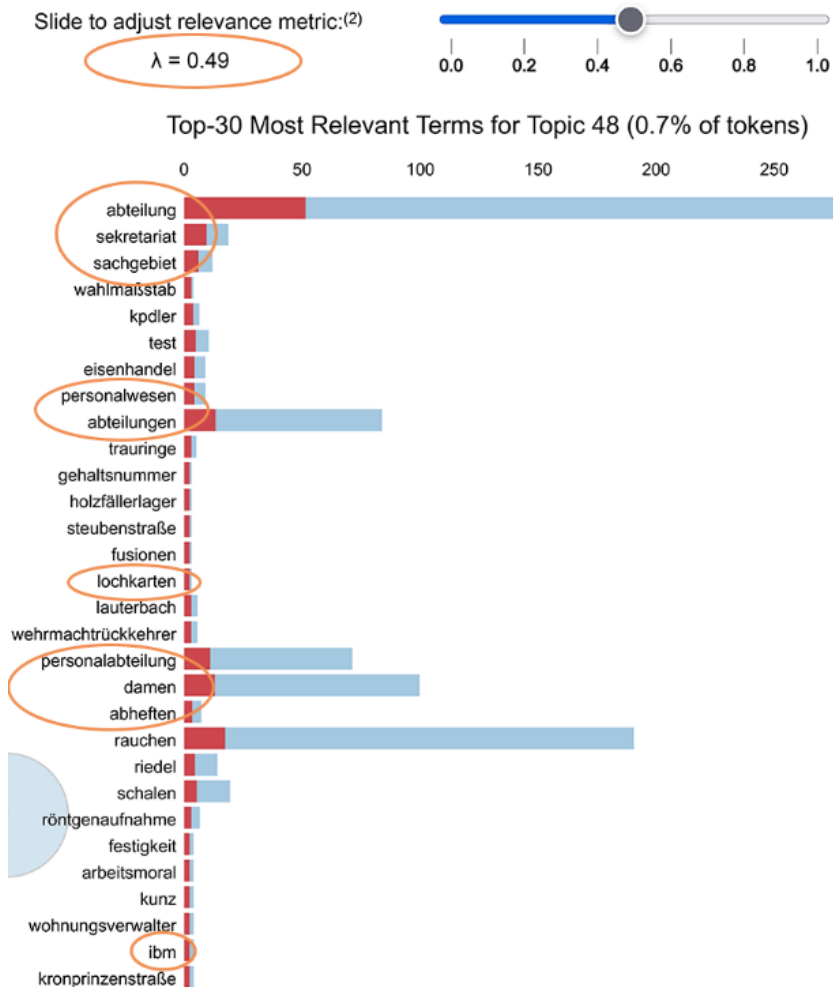


Abb. 4: Reduktion der relevance metric über den Lambda-Regler im Topic 48 (Quelle: eigene Darstellung).

Schaut man sich die Ergebnisse näher an, stellt man fest, dass die für Oral-History-Interviews typischen Stegreiferzählungen das Topic Modeling an seine Grenzen bringen. Durch schnelle Themenwechsel, Abschwefungen und dialektale Sprache finden sich in vielen Topics dieselben Begriffe in geringerer oder häufigerer Anzahl (ubiquitär sind etwa „Krieg“ oder die Stadt „Essen“, vgl. Abb. 2). Trotz dieses stetigen „Rauschens“ lassen sich in den unterschiedlichen Topics thematische Zuspitzungen erkennen. Folgt man beispielsweise dem Begriff „Abteilung“, der in der Übersicht der häufigsten Keywords zwischen „Krieg“ und „Essen“ liegt, gelangt man zu Topic 48, visualisiert als einer der blauen Kreise im Koordinatenfeld der grafischen Darstellung. Die

Zusammensetzung des Topics verweist auf den diffusen Charakter der lebensgeschichtlichen Interviews, gibt aber deutliche Hinweise auf die Darstellung innerbetrieblicher Erfahrungen (vgl. Abb. 3).

Es fällt auf, dass hier keine schwerindustriespezifischen Themen angerissen werden, sondern anhand der Begriffe „Personalabteilung“, „Büro“ und „Sekretariat“ die Verwaltungsarbeit in den Fokus rückt. Auffällig ist der explizite Verweis auf die „Damen“ und die „Sekretärin“, weshalb sich die Frage nach Geschlechterverhältnissen in der Ruhr-Industrie aufdrängt. Doch folgt man der begrifflichen Verästelung des Topics in die Tiefe, rückt ein weiterer Aspekt ins Blickfeld. Dazu wird die *relevance metric* in der interaktiven visuellen Darstellung heruntergeregt. Das bedeutet, dass ein Algorithmus auch die global weniger häufigen, aber mit dem Topic exklusiver verknüpften Begriffe in die Übersicht einbezieht. Der Ausschlag des Balkendiagramms gibt dabei das Verhältnis von Gesamtworthäufigkeit und Vorkommen des einzelnen Begriffs wieder (Sievert/Shirley 2014: 66 f.).

Hier liegt die besondere Qualität des Topic Modeling – insbesondere in Kombination mit einem Explorations-Tool wie *pyLDAvis*. So wird es möglich, systematisch nach statistischen Ausreißern zu suchen: Die Begriffe „Locharten“ und „IBM“ kommen im gesamten Sample verschwindend selten vor, sind aber für Topic 48 von Bedeutung. Der rote Balken, der das Vorkommen dieser seltenen Begriffe innerhalb des Topics misst, ist beinahe deckungsgleich mit dem blauen Balken, also dem Vorkommen des Begriffs im Verhältnis zum Gesamtsample. Das bedeutet, insgesamt ist der Begriff wenig repräsentativ für das Große und Ganze, für das Topic 48 ist er jedoch besonders charakteristisch, weil er fast ausschließlich in diesem und in keinen anderen Topics vertreten ist. Die semantische Evaluation bestätigt die Berechnungen: Zwischen den in Abbildung 3 eingekreisten Begriffen besteht ein direkter inhaltlicher Zusammenhang. Folgt man den Oberbegriffen mit Hilfe des Tools *pyLDAvis* in die Tiefen der Topics, gelangt man also vom Allgemeinen zum Besonderen. Dem induktiven Vorgehen der qualitativen Forschung wird damit ein algorithmengestützter Zugang an die Hand gegeben, um große unstrukturierte Quellensammlungen systematisch zu explorieren.

Im vorliegenden *Use Case* werden auf diese Weise frühe Automatisierungsprozesse in der Verwaltung sichtbar. Zunächst erscheint der Begriff „Lochkarten“, kurz darauf das Unternehmen „IBM“. Das sind Themen, mit denen man in einer Sammlung zu Kriegserfahrungen und Industriearbeit nicht unbedingt rechnet und die leicht übersehen werden können.

Tatsächlich ist die Erforschung zunehmender Automatisierung und Digitalisierung von Verwaltungsapparaten – insbesondere die weibliche Perspektive, die die Begriffe „Damen“ und „Sekretärinnen“ im Topic 48 verheißen – ein Desiderat der zeitgeschichtlichen Forschung. So spielen auch in den Aufsätzen von Margot Schmidt, die einen Großteil der weiblichen Angestellten im LUSIR-Projekt interviewt hat, die Beschäftigungsfelder, nicht jedoch die spezifischen Aufgaben im Arbeitsalltag der Frauen eine Rolle. In ihren damaligen Untersuchungen zur Stellung der Frau in den Ruhrgebietsindustrien zwischen 1930 und 1980 gerieten die Modernisierungsprozesse der Unternehmensverwaltungen zur Marginalie (vgl. Schmidt 1983a: 159) oder sogar ganz aus dem Blick (vgl. Schmidt 1983b).

Vom *distant reading* zum *close reading* I: Zukunftsängste

Mit Hilfe des Topic Modeling konnten in nur wenigen Schritten historisch bedeutsame Informationen ausfindig gemacht werden, die unter der Oberfläche der LUSIR-Sammlung schlummerten. Im nächsten Schritt wurden die mit den Suchergebnissen verknüpften Passagen der Interviewtranskripte einer klassischen, analogen, hermeneutischen Lektüre, einem *close reading*, unterzogen. Ein interaktiver Rückschluss per Mausklick von *pyLDavis* in die verknüpfte Interviewpassage ist nicht möglich. Aber über die Dokument-Topic-Zuordnungen in *Gensim*, womit die Berechnung des Topic Modeling durchgeführt wurde, können die mit Topic 48 assoziierten Textabschnitte oder die gesamten Interviews, aus denen die Abschnitte stammen, ausgegeben werden.

Die Themen Digitalisierung und Automatisierung von Verwaltungsprozessen kommen sowohl bei Männern als auch bei Frauen des untersuchten Samples vor. Das ermöglicht die Gegenüberstellung der Erfahrung und Bewertung von Modernisierungsprozessen nach Geschlechtern. Der 1931 geborene Klaus Gerber (Pseudonym), Sohn eines Kranführers und ausgebildeter Chemotechniker, berichtet zunächst von den öffentlichen Gehaltsverhandlungen der 1960er Jahre, als das Gespräch auf Veränderungen in der Arbeitswelt einschwenkt. Auf Nachfrage ergänzt er, dass heutzutage – also in den 1980er Jahren, zur Zeit des Interviews – Arbeitsplatzsicherung Priorität habe:

K.G.: Ja, sicherlich ist die Arbeitsplatzsicherung heute an erster Stelle gerückt und vor allen Dingen die Besitzstandsicherung im Hinblick auch auf die erreichte Lohngruppe. Denn zurzeit, ich sehe das bei uns in der Widia so, wird sehr stark der Einsatz von nummerisch gesteuerten Werkzeugmaschinen betrieben.

I(interviewer): Nummerisch?

K.G.: Nummerisch gesteuerte. NC oder heute gibt es ja schon die Computer, die CNC-Maschinen auch schon. Und da besteht für viele Kollegen, die vorher hochwertige Facharbeiten gemacht haben, wo die herkömmlichen Dreh-, Fräsbänke und so weiter, entfallen, die Angst, dass sie aufgrund der Maschine, die einmal erworbene Lohngruppe verlieren könnten. Weil eben die Ansprüche der Automaten, wenn ich das mal in Anführungsstriche setzen soll, an die Fachkenntnisse nach Auffassung der Arbeitgeber nicht mehr so hoch sind. Das ist zurzeit die Angst der Kollegen, die also davon betroffen sind in den Betrieben bei dieser Umstellung. Hat erst jetzt begonnen. Die wird sich in den nächsten Jahren noch verstärkt fortsetzen. Und eines Tages wird das Problem beherrschend sein

(Interview mit Klaus Gerber (Pseudonym), geboren 1931, Interviewer: Ulrich Herbert, Datum: 23.2.1982. Archiv „Deutsches Gedächtnis“ im Institut für Geschichte und Biographie, FernUniversität in Hagen, Bestand LUSIR).

Als Chemotechniker bei der Widia, einem Subunternehmen der Krupp-Gruppe, hatte Gerber durch diese Verfahren selbst nichts zu befürchten, als Gewerkschaftsmitglied verfügte er allerdings über die nötige Sensibilität für die Ängste seiner Kollegen vor der Computerisierung. Dabei wurde der zunächst scheinbar offensichtliche Verlust von Arbeitsplätzen durch den Einsatz von NC/CNC- oder CAD-Technologien bereits in den 1980er Jahren differenziert diskutiert und in den letzten Jahren deutlich relativiert. Es

wurden beispielsweise weniger ausgebildete Dreher benötigt, während der Bedarf an speziell ausgebildeten Fachkräften für die computergesteuerten Maschinen rapide anstieg. Das führte wegen unzutreffender Prognosen und Fehlentscheidungen im Ruhrgebiet sogar kurzzeitig zu einem Fachkräftemangel, dem durch Arbeitsmigration begegnet werden musste (Raphael 2019: 58 ff.).⁸ Gerber macht nicht deutlich, ob er sich um die Perspektiven qualifizierter Arbeiter sorgt oder ob er in der Computerisierung ein gesellschaftsübergreifendes „beherrschendes Problem“ sieht. Zumindest war die Bedrohung der klassischen Arbeitsplätze durch Computer seit der Entstehung der Großrechner, der sogenannten *Mainframes*, in den 1950er Jahren ein gängiges Narrativ in der deutschen Öffentlichkeit (Bösch 2018: 21; Schlombs 2019: 222 f.).⁹

Auch wenn aktuelle Studien zeigen, dass Automatisierung und Digitalisierung nicht grundsätzlich zu Arbeitsplatzabbau oder gar Entlassungen führen, prognostiziert der 1915 geborene Walter Paulsen (Pseudonym), Sohn eines Facharbeiters, selbst ausgebildeter Rechtsanwalts helfer und Personaler bei Krupp, im Interview 1981 schwere Zeiten für Angestellte in der Unternehmensverwaltung. Obwohl er zum Zeitpunkt des Interviews schon wesentlich älter war als Gerber, zeigt sich Paulsen eher fasziniert von der Entstehung einer digitalisierten Verwaltung. Dazu bedient er sich eines weiteren Stereotyps: dem Vergleich von menschlicher und maschineller Intelligenz (vgl. Schlombs 2019: 222 f.).

W.P.: Wer weiß, ob in meinem Beruf, ob man heute, wenn man das Verwaltungsfach einschlägt so Personal, ja, ob man das sein ganzes Leben durch machen kann.

I.: Hat sich ja auch der Verwaltungsberuf hat sich ja völlig verändert.

W.P.: Ja sicher.

I.: Das ist ja alles EDV jetzt. [...] Das sind ja wirklich ganze, ganze Abteilungen sind da ja stillgelegt worden.

W.P.: Klar sicher. Einer muss noch das weiterführen, das ist klar, ne. Und, äh, sicher, man kann natürlich den Computer, ich weiß es nicht, ob das noch kommt, dass der Computer nun fragt, äh, bei Einstellungsgesprächen und so weiter, ob das möglich ist, das weiß ich nicht. Vielleicht kommt das auch noch ja?

(Interview mit Walter Paulsen (Pseudonym), geboren 1915, Interviewer: Ulrich Herbert, Datum: 4.7.1981. Archiv „Deutsches Gedächtnis“ im Institut für Geschichte und Biographie, FernUniversität in Hagen, Bestand LUSIR).

Paulsen ordnet die EDV kurz darauf selbst technikgeschichtlich ein:

W.P.: Ja. Man brauchte dann nicht mehr diese manuelle Gehaltsabrechnung machen, ne, das wurde alles eingegeben in die EDV dann, ne. [...] Damals nannte sich das Hollerith, ja? Hollerithabteilung, ja. Die wurde denn ausgebaut, ja. [...] Da haben wir Gehaltskarten gehabt, ja, und die Gehaltskarten ... Früher wurde ja alles manuell, da hatten wir ungefähr zehn Gehaltsbuchhalter, ja, und auf einmal hatten wir nur noch drei oder vier, ja?

⁸ Vgl. zur gesellschaftlichen Debatte: Kemmer (1985).

⁹ Vgl. Titel „Die Roboter sind unter uns“ und Leitartikel „Die Revolution der Roboter“, in: Der Spiegel, Nr. 31/1955. Online unter: <https://www.spiegel.de/spiegel/print/index-1955-31.html>, <https://mazin.spiegel.de/EpubDelivery/spiegel/pdf/31970830> (24.1.2021).

I.: Aha. Ach das warn Lochkartensystem das [...] Hm. Und dann Umstellung auf EDV.

W.P.: Ja, das, äh, in den 60er Jahren denn nachher EDV, ne. Das hatte natürlich auch Schwierigkeiten, wegen der Geheimhaltung, ja, aber das hat sich mit der Zeit doch durchgesetzt. [...] Denn ich weiß, Neunzehnhundert ..., äh, die leitenden Angestellten, ja, die wurden zuerst noch alle manuell abgerechnet, ja? Und das hab' ich auch für die Direktoren denn auch noch selbst gemacht dann, ne. Aber eines Tages dann hatte man denn, dann hörte man auch, was der verdient, was der verdient, dann stand's am schwarzen Brett, ne, oder in der „Kapital“ hat ja, oder „D-Mark Zeitung“, ja die haben ja die Gehälter alle veröffentlicht dann, ne. [...] Und dann ist man dazu übergegangen, dass man auch die leitenden Herren sogar in die EDV reingetan hat, ja. Hat man eben keine Nummer, den Namen nicht geschrieben, nur die Nummer, ja, ne, und dann war das ja. (ebd.).

Datenschutzbedenken wurden in den frühen 1980er Jahren durchaus schon öffentlich diskutiert (vgl. N. N. 1983). Das Hollerithverfahren hatte die Grundlage für die EDV gelegt und war im späten 19. Jahrhundert vom Namensgeber Herman Hollerith, Nachfahre deutscher Amerikauswanderer, erfunden worden. Er entwickelte eine Tabelliermaschine, die Daten in Lochkarten speichern konnte (Driessen 1987: 43) – über den Begriff „Lochkarte“ ist die Aufmerksamkeit mit Hilfe des Topic Modeling auf diese für die Geschichte der Digitalisierung so aufschlussreiche Interviewpassage gelenkt worden.

Vom *distant reading* zum *close reading* II: Holleriths Vermächtnis

Bis zur Entstehung des elektronischen Computers wurden Hollerithmaschinen flächendeckend in der Verwaltungsarbeit eingesetzt und gelten einigen Technikhistorikern sogar als wichtiger Impuls für die Erforschung künstlicher Intelligenz (Breslauer 2013: 105; Nilsson 2010: 23). Seit dem Zweiten Weltkrieg wurden zunehmend Frauen an Lochkartensystemen beschäftigt, wie man auch den LUSIR-Interviews entnehmen kann (Schmidt 1983a: 134). Diese zunächst monotone und stupide Arbeit wurde für viele Frauen zum Türöffner in die Verwaltungsarbeit, was auch im Folgenden noch näher beschrieben wird (Jung 2015: 101 ff.; Meier 1987: 49). Irmgard Schonfeld (Pseudonym), Tochter eines Drehers und einer Köchin, wurde 1939 mit 18 Jahren ohne vorherige Berufsausbildung bei Thyssen an der Hollerithmaschine angelernt:

[...] Und da sind wir dann angelernt worden und, äh, so kamen dann nachher immer mehr Frauen dazu, nicht. Wir waren damals anfangs dann sechs Frauen, ne. So wurde dann in der Lochkartenabteilung, wie sie sich noch nannte, äh, kamen dann langsam Frauen rein. Dann habe ich gelocht, geprüft, ich habe das richtig von der Pike auf gemacht und bin dann, weil ich Interesse zeigte und mir das auch lag, nachher in den Maschinenraum gekommen und hab Schalten gelernt, hab Lehrgänge besucht in Berlin.

(Interview mit Irmgard Schonfeld (Pseudonym), geboren 1921, Interviewerin: Margot Schmidt, Datum: 9.2.1982. Archiv „Deutsches Gedächtnis“ im Institut für Geschichte und Biographie, FernUniversität in Hagen, Bestand LUSIR).

Nach einem kurzen Intermezzo in der Textilindustrie kehrte sie zurück nach Duisburg und in die Metallindustrie:

[...] Und dann hörte ich, dass auf der niederrheinischen Hütte, die ja auch zur Thyssenhütte gehörte, aber selbständig war, dass da eine Lochkartenabteilung aufgemacht werden sollte, und da habe ich mich ganz schnell beworben, ich kannte ja noch die Herren alle, und dann bin ich dahin gekommen als Leiterin vom Lochraum, ne (ebd.).

Nach sechs Jahren kam sie als Maschinenraumaufseherin wieder zum Mutterkonzern Thyssen. Sie beschreibt die technische Entwicklung und die Vergrößerung ihrer Abteilung:

[...] da habe ich so, und auch von Maschinen, Sortiermaschinen, Tabelliermaschinen, Rechenlocher, habe ich so das alles mitgemacht. Früher gab es nur IBM, habe ich die ganzen Stufen mitgemacht, immer wieder 'ne größere Maschine. [...] Immer wieder neu lernen, und so habe ich viele, sehr viele Lehrgänge mitgemacht auf Niederrhein und auch nachher hier noch. Joah, und Leiterin bin ich geblieben von Anfang an. Hab hier unheimlich viele Damen durchgehend, durchgehend gehabt, ich hatte schon, meine Höchstzahl war 96 Damen, ne. Das sind viel, 96, das war ein Wahnsinn (ebd.).

Eine Ausbildung für das Bedienen der Lochkartenmaschinen war nicht vorgesehen, und gemäß der Losung von IBM – „jeder kann lochen, prüfen und tabellieren“, wie Schonfeld erinnert – habe man „früher wahllos die Damen angenommen. Wer sich hier vorgestellt hat, hat nen kleinen Test gemacht, und wenn er gut war, ist er angenommen worden [...]“. Doch mit weiteren technologischen Entwicklungsschritten und der zunehmenden Professionalisierung von Verwaltungsvorgängen wuchsen auch die Anforderungen an die Bewerberinnen (vgl. Schmidt 1983a: 140 f.; Schmidt 1983b: 214). Ab Anfang der 1970er Jahre absolvierten die jungen Frauen eine einjährige Ausbildung, die Maschineschreiben, Mathematik und Rechnungswesen umfasste, bevor sie ihre Arbeit in der Abteilung von Irmgard Schonfeld antraten.¹⁰ Technologischer Fortschritt bedeutete allerdings auch Rationalisierung, wie in den beiden Interviews mit den männlichen Angestellten angedeutet wurde. Während Schonfelds Stelle sicher war, verringerte sich die Belegschaft ihrer Abteilung im Laufe der 1970er Jahre zusehends:

Äh, ich habe jetzt noch, eins, zwei, eins zwei, drei vier, fünf Damen, das sind dann aber schon ältere, aber was heißt älter, nicht über vierzig, ne, die keine Schul..., unsere Schulausbildung haben, ja. Die anderen sind alle aus diesen Klassen, ja. Die ersten sind natürlich schon weg, ne. Aber, äh, die anderen sind noch hier, ne. Und inzwischen sind wir, sind wir ja von IBM fast ab, wir haben Nixdorf, das ist ja Platte und Band, kennen Sie, ne, und so ist von den alten Maschinen nichts mehr da. Es sind noch 'n paar so ein paar kleine Arbeiten werden da noch drauf gemacht, und das ist so ja alles (ebd.).

¹⁰ Vgl. zur frühen Diskussion über Anforderungen bei Ausbildung für das durch EDV unterstützte Büro: Schmid 1971.

Der Aufstieg Irmgard Schonfelds ist gewiss nicht der Normalfall, zeigt aber die Möglichkeiten auf, die sich Frauen in den Nachkriegsjahren in den von Männern dominierten und patriarchal geleiteten Industrieunternehmen des Ruhrgebiets boten (vgl. Schmidt 1983b: 203 ff.). Der Mangel an Männern in den Kriegsjahren mag ein wichtiger Faktor für Schonfelds Aufstieg gewesen sein, doch in der unmittelbaren Nachkriegszeit versuchte man, die Sozialstruktur *ante bellum* wieder herzustellen und stellte zunächst keine weiteren Frauen in der Industrie ein (Schmidt 1983a: 156). Schonfeld wird sehr deutlich, wenn Geschlechterfragen zu Sprache kommen. Sie erzählt von ihrem gleich gut qualifizierten Kollegen, der gemeinsam mit ihr bei Thyssen angefangen hatte, jedoch bessere Aufstiegsmöglichkeiten verwirklichen konnte. Auf die Frage, ob das geschlechterbedingt gewesen sei, antwortet sie:

I.S.: Ohne weiteres. Ja. Ja. Denn als ich anfing, da waren ja die Frauen nichts, ne. Das haben Sie ja da schon dann von gehört, der höchste Chef damals, der wollte ja keine Frauen beschäftigt haben, hat sich ja unheimlich [...] gesträubt, Frauen zu beschäftigen. Und wir sind dann wie so Wundertiere durch die Verwaltung gegangen, ja. Überall wurden wir angestarrt, weil die das ja nicht kannten, auf einmal Frauen im Büro, ne.

I: Ist ja wirklich komisch. Und als Sie dann angefangen haben mit ihrem Kollegen hier zu arbeiten, das waren am Anfang, waren das ja alles noch Männer?

I.S.: Waren alles Männer, ja.

I: Wie war so das Arbeitsklima?

I.S.: Ich kann nicht klagen, ne. Die haben uns nachher, wie man so schön sagt, für voll genommen, wir mussten ja, äh, von denen wiederum lernen, ja.

I: Mhm. Sie haben Ihnen das beigebracht?

I.S.: Die haben uns das beigebracht. Ja. Ja. Schalten, ich weiß nicht, ob Sie sowas mal gesehen haben. Dann gab's so Schalttafeln, gab's 'ne ganz alte Maschine, die, früher war's noch alles einfach, da wurden nur die Posten addiert, und äh, subtrahieren konnte man da dran, aber man musste alles Schalten mit, mit Kabeln und mit zwei Steckern, ne. Dann hatte die Maschine, das war eine alte D11, solche große Klappe, die wurde vorne aufgemacht und dann hat man auf einem Fußbänkchen gesessen, und dann hat man, sagen wir mal, Gehaltsnummer, Abteilung und Betrag nehmen wir jetzt mal nur, da mussten sie alles schalten. Die Lochkarte hat ja achtzig Stellen, dann stand meinetwegen die Gehaltsnummer von eins bis fünf, die Abteilung von sechs bis acht und der Betrag von neun bis zehn, bis, bis zwanzig oder so. Das mussten sie alles ... und das gleiche, diese gleiche Stellenzahl von eins bis achtzig war auch in der Maschine, und das mussten sie dann schalten, richtig stecken, wo das stehen sollte, und die, äh, Walze von der Maschine, die hatte auch achtzig Stellen. Und wenn ich das jetzt, meinetwegen meine Gehaltsnummer, die von eins bis fünf stand, auf der Walze dann von zehn bis fünfzehn geschaltet hatte, dann musste ich das von eins nach zehn von zwei nach elf usw. schalten, ne. Und dann kamen nachher die vier einundzwanzig, da waren kleinere Schalttafeln schon, immer moderner und, äh, auch immer schwieriger, weil immer mehr verlangt wurde. Heute ist das ja schon ins Unermessliche, ne (ebd.).

Während Frauen den Männern in diesem Arbeitsbereich in der frühen Nachkriegszeit zahlenmäßig unterlegen waren und von diesen eingewiesen und angeleitet wurden, kehrte sich der Trend im Laufe der Jahrzehnte um, wie Schonfeld darstellt:

I.S.: Ja, der Arbeitsplatz ist, hat sich mit den Jahren immer mehr modernisiert und ist immer, ähm, was eine Locherin oder heute, äh, eine Datentypistin anbelangt, ganz ins weibliche übergegangen, ne.

I.: Die Männer sind dann irgendwie so verschwunden.

I.S.: Ja, die sind verschwunden. Ja. Die waren dann nachher im Maschinenraum, ne zuerst, Maschinenraum sagte man dazu, zu diesen Räumen da, ne. Und äh, jetzt inzwischen sind das Sachbearbeiter geworden, ne. [...]

I.: Und so, Konkurrenz untereinander?

I.S.: [...] Die war anfangs. Eben, äh, große Neider, ja. [...] Deswegen habe ich auch sehr wenig, äh, arbeitsmäßig sehr wenig Kontakt mit Frauen gehabt, ja. Ich konnte sehr gut mit Männern zusammenarbeiten, hab auch jahrelang im Maschinenraum nur mit Männern gearbeitet, weil ich Tabelliererin war oder wie sich das zurzeit nannte, ne, und kann besser mit Männern zusammenarbeiten als mit Frauen, ja (ebd.).

Irmgard Schonfelds Erzählung gewährt einen direkten Einblick in die Reziprozität von Frauenarbeit und technologischer Entwicklung in den Verwaltungen westdeutscher Industrieunternehmen der Nachkriegszeit. Auch wenn die Aufgaben keine intellektuelle Herausforderung darstellten, erforderten die neu zu erlernenden Praktiken Konzentrationsfähigkeit und den Willen, ein neues, technisch geprägtes Betätigungsfeld zu erschließen. Frau Mittler (Pseudonym), Tochter eines Ingenieurs und einer Schneiderin, gehörte zur Generation, die den Pionierinnen um Schonfeld bei Thyssen folgte. 1933 geboren, begann Mittler im Alter von 16 Jahren und nach Abschluss einer kaufmännischen Lehre ihre Arbeit bei Thyssen. Aus Sozialstruktur und technologischem Fortschritt resultierende Probleme sieht sie allerdings weniger bei geschlechterspezifischen als viel mehr bei Generationsfragen:

F.M.: Ja, also mit dem Geschlecht hing das an und für sich nicht zusammen, das ist ganz klar, die, die hatten ihre Methode, und dann wurde was anderes angeordnet, dass man's anders macht, und dann fing das auch mit der Datenverarbeitung an, und da waren vor allen Dingen ältere Leute, die waren ein bisschen überfordert, das ist ja ganz klar, dann, aber richtig Datenverarbeitung war dann nachher nicht so doll. Und da habe ich, da waren viele Ältere, die dann nicht mehr mitmachen konnten, die schon sechzig oder noch älter waren, die konnten sich daran nicht gewöhnen. [...]

I.: Sind die dann irgendwie umgesetzt worden, oder?

F.M.: Nö, die blieben an und für sich in der Abteilung. Wir hatten dann so einen alten Herren, ja, der konnte sich nicht, der hat der Maschine nicht getraut, oder dann haben wir hier auch noch vor Jahren gehabt, dass jemand, äh, die Steuern noch errechnete, was die Maschine ja brachte, ne? Und die haben der Maschine nicht getraut, die haben auch nach wie vor die Beträge auf die Karte geschrieben, weil die dachten, wir, mit Listen können wir nichts anfangen, wir müssen

die Karte da ziehen können, wo von dem Mann dann zwölfmal hintereinander steht [...].

I.: Wenn man das Jahrzehnte so gemacht hat, dann ist das auch schwierig.

F.M.: Ja, ist auch schwierig, da weiß man auch so bei Kleinigkeiten, wenn sich irgendwie was ändert oder so, man ist anfangs unsicher, und wenn man älter wird, die älteren Leute, der Wohnort, diesen Kollegen, von dem ich sprach, der sich da nicht dran gewöhnen konnte, der kam zum Beispiel aus F., der hatte sechs Stunden Fahrt pro Tag, der hatte da Eigentum, der konnte ja auch nicht umziehen, war ganz schlimm, ja. Und dann sollte sich der da noch mit diesem neuen Kram da zu Recht finden, das war ja ganz unmöglich. Ja, und dann kam da die Alters ..., der Altersstarrsinn kommt so ein bisschen dazu, ist ja klar, dass da zwischen jungen Leuten, die fortschrittlicher waren, und das war auch so, wenn, wenn Abteilungen zusammen gelegt werden, das ist immer so. Das hat ja mit Frau oder Mann nichts zu tun, bisschen Konkurrenz, jeder will der erste sein, das ist auch ganz klar, ne?

(Interview mit Frau Mittler (Pseudonym), geboren 1933, Interviewerin: Margot Schmidt, Datum: 5.7.1982. Archiv „Deutsches Gedächtnis“ im Institut für Geschichte und Biographie, FernUniversität in Hagen, Bestand LUSIR).

Wenn man die persönlichen Erfahrungen der frühen Automatisierung und der späteren Digitalisierung vergleicht, werden bis heute gängige Narrative computerbedingter Zukunftssängste und -hoffnungen sichtbar. Interessant ist, wie diese technologischen Modernisierungsprozesse Frauen ermöglicht haben, *Future Skills* zu erlernen, die zum Einzug und Aufstieg von Frauen in der Verwaltung beigetragen haben. Eine kontrovers diskutierte These ist, dass insbesondere Frauen vom erstarkenden tertiären Sektor in den 1970er Jahren profitiert haben (Jung 2015: 101). Zieht man die von Yong Suk Jung zusammengestellten Arbeitsmarktstatistiken heran, wuchs die Beschäftigung von Frauen im Dienstleistungssektor zwischen 1985 und 2002 kontinuierlich an. 2002 arbeiteten gut zwei Drittel der erwerbstätigen Frauen im Dienstleistungsbereich, wohingegen nur ein Drittel aller männlichen Beschäftigten im tertiären Sektor beschäftigt war (ebd.: 107 f.). Differenziert man weiter, waren 2002 ein Viertel aller erwerbstätigen Frauen im Büro tätig (ebd.: 108). Damit korrespondiert, dass die Frauen im vorliegenden Sample die düsteren Aussichten der Männer bezüglich der Automatisierung nicht teilten und stattdessen die geschlechtsspezifischen Implikationen der technologischen Entwicklung reflektierten. Für sie manifestierten sich im Computer Karriereperspektiven und Zukunftshoffnungen.

Am Anfang standen stupide und monotone Aufgaben für ungelernte Arbeiterinnen, die sich für viele Frauen auch am Computer fortsetzen sollten – so ließ die ZEIT 1974 verlauten: „Taylor schleicht sich ins Büro“ (Jungblut 1974). Doch die wachsende EDV verlangte zunehmend nach qualifizierten Frauen, die sich so Schlüsselpositionen in den unteren und mittleren Verwaltungsebenen vor allem kleiner und mittelgroßer Unternehmen sichern konnten (Eder/Woll 1987: 41; Meier 1987: 49 f.). Mitunter waren die Sekretärinnen sogar früher mit der Digitalisierung konfrontiert als die Männer in den Führungspositionen selbst (Holtgreve 1997: 252; Wagner 1986: 54). Die Ironie daran ist, dass Frauen somit über Jahrzehnte hinweg die Führungspositionen der Männer si-

cherten, indem sie den zunehmend computergestützten „Vorzimmeralltag“ oder komplexe Verwaltungsvorgänge meisterten, ohne selbst in bedeutendem Maße die Karriereleiter weiter hinaufzusteigen (Eder/Woll 1987: 41; Johanson/Taylor 1986: 81).

Fazit

Wie die kleine explorative Studie gezeigt hat, kann diese Entwicklung mit einer Sekundäranalyse bestehender Interviewsammlungen untersucht werden. Durch die Anwendung des Topic Modeling wurde das vorliegende Interview-Korpus statistisch vermessen und inhaltlich aufgeschlüsselt. Innerhalb der Ergebnisse konnte induktiv und explorativ entlang häufiger Keywords immer tiefer in die Verästelungen verwandter Begriffe vorgedrungen werden, um die besonderen Randphänomene allgemeiner Themen systematisch sichtbar zu machen. In einer Interviewsammlung zu Faschismus- und Industriearbeitserfahrungen kamen so Erzählungen über die Frühgeschichte der Automatisierung und Digitalisierung von Unternehmensverwaltungen zum Vorschein, in der Frauen eine signifikante Rolle gespielt haben. Im demonstrierten Fall konnte das Topic Modeling seine Stärken als hilfreiche explorative Methode zur Erschließung unstrukturierter Interviewsammlungen in der historischen Biographieforschung unter Beweis stellen.¹¹

LITERATUR

- Blei, David M. (2012): Topic Modeling and Digital Humanities, in: *Journal of Digital Humanities*, 2, Number 1. Online: <http://journalofdigitalhumanities.org/2-1/topic-modeling-and-digital-humanities-by-david-m-blei/> (25.11.2021).
- Bösch, Frank (2018): Wege in die digitale Gesellschaft, Computer als Gegenstand der Zeitgeschichtsforschung, in: Ders. (Hg.): *Wege in die digitale Gesellschaft, Computernutzung in der Bundesrepublik 1955-1990, Geschichte der Gegenwart*, Bd. 20, Göttingen, 7-36.
- Breslauer, Burkhardt (2013): Von der Datenverarbeitung zur Informationstechnologie – „Mythos IBM“, Globale Netzwerke und lokales Engagement in der Wirtschaft am Beispiel der IBM Deutschland GmbH, in: *Zeitschrift für Weltgeschichte*, 14, Heft 1, 101-121. https://doi.org/10.3726/84540_101
- Driessen, Tilman (1987): Von Hollerith zu IBM, Zur Frühgeschichte der Datenverarbeitungstechnik von 1880 bis 1970 aus wirtschaftswissenschaftlicher Sicht, *Wirtschafts- und Rechtsgeschichte*, Bd. 5, Köln.
- Eder, Harriet und Ellen Woll (1987): Arbeit und Geschlecht, in: Gerhard Brosius und Frigga Haug (Hg.): *Frauen/Männer/Computer, EDV im Büro, Empirische Untersuchungen, Argument/Sonderband, AS 151*, herausgegeben von der Hochschule für Wirtschaft und Politik, Berlin, 40-42.
- Fechner, Martin und Andreas Weiß (2017): Einsatz von Topic Modeling in den Geschichtswissenschaften: Wissensbestände des 19. Jahrhunderts, in: *Zeitschrift für digitale Geisteswissenschaften*, Heft 2. http://doi.org/10.17175/2017_005.
- Graham, Shawn, Ian Milligan und Scott Weingart (2016): *Exploring Big Historical Data, The Historian's Macroscope*, London. <https://doi.org/10.1142/p981>
- Holtgrewe, Ursula (1997): *Frauen zwischen Zuarbeit und Eigensinn, Der EDV-Einzug in Kleinstbetriebe und die Veränderungen weiblicher Assistenzarbeit*, Berlin.

¹¹ Vgl. <https://tedunderwood.com/2012/04/07/topic-modeling-made-just-simple-enough/> (13.8.2021).

- Johanson, Bonnie McD und James Taylor (1986): Informationsverarbeitung und Karrieremöglichkeiten für Frauen, in: Paul Kolm und Ina Wagner (Hg.): *Frauen, Arbeit und Computerisierung*, Schriftenreihe des Institutes für Berufsbezogene Erwachsenenbildung an der Johannes-Kepler-Universität, 5, Nr. 2, Linz, 60-85.
- Jung, Yong Suk (2015): Strukturwandel im sozialen Feld, Bergarbeiterfamilien im Ruhrgebiet 1945 bis 2000, Veröffentlichungen des Instituts für Soziale Bewegungen, Schriftenreihe A, Darstellungen, Bd. 54, Essen.
- Jungblut, Michael (1974): Taylor schleicht sich ins Büro, in: *Die ZEIT*, Nr. 38. Online: <https://www.zeit.de/1974/38/taylor-schleicht-sich-ins-buero> (24.1.2021).
- Kemmer, Heinz-Günter (1985): Mangel im Überfluß, in: *Die ZEIT*, Nr. 31. Online: <https://www.zeit.de/1985/31/mangel-im-ueberfluss> (24.1.2021).
- Leh, Almut, Joachim Köhler, Michael Gref und Nikolaus P. Himmelmann (2018): Speech analytics in research based on qualitative interviews. Experiences from KA3, in: *VIEW, Special Issue: Audiovisual Data in Digital Humanities*, 7, Issue 14, 138-149. <http://doi.org/10.18146/2213-0969.2018.jethc158>.
- Meier, Sigrid (1987): Neue Arbeitsteilungen?, in: Gerhard Brosius und Frigga Haug: *Frauen/Männer/Computer, EDV im Büro, Empirische Untersuchungen, Argument/Sonderband, AS 151*, herausgegeben von der Hochschule für Wirtschaft und Politik, Berlin, 47-53.
- N. N. (1983): Erfäßt, registriert, entmündigt, Die Volkszählung bringt die Bürger auf die Barrikaden: Ist die Datengier der Bürokraten noch zu zügeln?, in: *Die ZEIT*, Nr. 12. Online: <https://www.zeit.de/1983/12/erfasst-registriert-entmuendigt> (24.01.2021).
- Niethammer, Lutz (Hg.) (1983a): „Die Jahre weiß man nicht, wo man die heute hinsetzen soll“, *Faschismuserfahrungen im Ruhrgebiet, Lebensgeschichte und Sozialkultur im Ruhrgebiet 1930-1960*, Bd. 1, Berlin, Bonn.
- Niethammer, Lutz (Hg.) (1983b): „Hinterher merkt man, daß es richtig war, daß es schiefgegangen ist“, *Nachkriegserfahrungen im Ruhrgebiet, Lebensgeschichte und Sozialkultur im Ruhrgebiet 1930-1960*, Bd. 2, Berlin, Bonn.
- Niethammer, Lutz und Alexander von Plato (Hg.) (1985): „Wir kriegen jetzt andere Zeiten“, *Auf der Suche nach der Erfahrung des Volkes in nachfaschistischen Ländern, Lebensgeschichte und Sozialkultur im Ruhrgebiet 1930-1945*, Bd. 3, Berlin, Bonn.
- Nilsson, Nils J. (2010): *The Quest for Artificial Intelligence, A History of Ideas and Achievements*, New York. <https://doi.org/10.1017/CBO9780511819346>
- Raphael, Lutz (2019): *Jenseits von Kohle und Stahl, Eine Gesellschaftsgeschichte Westeuropas nach dem Boom, Frankfurter Adorno-Vorlesungen 2018*, Berlin.
- Schlombs, Corinna (2019): *Productivity Machines, German Appropriations of American Technology from Mass Production to Computer Automation, History of Computing*, Cambridge (Massachusetts), London. <https://doi.org/10.7551/mitpress/11864.001.0001>
- Schmid, Klaus-Peter (1971) Neue Chancen im Büro, Arbeitskräftemangel zwingt zur Rationalisierung, in: *Die ZEIT*, Nr. 9. Online: <https://www.zeit.de/1971/09/neue-chance-im-buero> (24.01.2021).
- Schmidt, Margot (1983a): Krieg der Männer – Chance der Frauen?, in: Lutz Niethammer (Hg.), „Die Jahre weiß man nicht, wo man die heute hinsetzen soll“, *Faschismuserfahrungen im Ruhrgebiet, Lebensgeschichte und Sozialkultur im Ruhrgebiet 1930-1960*, Bd. 1, Berlin, Bonn. 133-162.
- Schmidt, Margot (1983b): Im Vorzimmer, in: Lutz Niethammer (Hg.), „Hinterher merkt man, daß es richtig war, daß es schiefgegangen ist“, *Nachkriegserfahrungen im Ruhrgebiet, Lebensgeschichte und Sozialkultur im Ruhrgebiet 1930-1960*, Bd. 2, Berlin, Bonn, 191-232.
- Sievert, Carson und Kenneth E. Shirley (2014): LDAvis: A method for visualizing and interpreting topics, in: *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, Baltimore, Maryland, USA, 63-70. <https://doi.org/10.3115/v1/W14-3110>

Spiegel 31/1955. Online: <https://www.spiegel.de/spiegel/print/index-1955-31.html>, <https://magazin.spiegel.de/EpubDelivery/spiegel/pdf/31970830> (24.1.2021).

Wagner, Ina (1986): Frauenarbeit im automatisierten Büro, in: Paul Kolm und Ina Wagner (Hg.): Frauen, Arbeit und Computerisierung, Schriftenreihe des Institutes für Berufsbezogene Erwachsenenbildung an der Johannes-Kepler-Universität, 5, Nr. 2, Linz, 43-59.

QUELLEN

Interview mit Klaus Gerber (Pseudonym), geboren 1931, Interviewer: Ulrich Herbert, Datum: 23.2.1982.

Interview mit Frau Mittler (Pseudonym), geboren 1933, Interviewerin: Margot Schmidt, Datum: 5.7.1982.

Interview mit Walter Paulsen (Pseudonym), geboren 1915, Interviewer: Ulrich Herbert, Datum: 4.7.1981.

Interview mit Irmgard Schonfeld (Pseudonym), geboren 1921, Interviewerin: Margot Schmidt, Datum: 9.2.1982.

Archiv „Deutsches Gedächtnis“ im Institut für Geschichte und Biographie, FernUniversität in Hagen, Bestand LUSIR

Zusammenfassung

Dieser Aufsatz dokumentiert die computergestützte Sekundäranalyse des Interviewbestands aus dem Oral History-Projekt Lebensgeschichte und Sozialkultur im Ruhrgebiet. Durch Anwendung des Text Mining-Verfahrens *Topic Modeling* wurde das Interviewkorpus statistisch vermessen und anschließend entlang inhaltsbeschreibender Keywords explorativ untersucht. Auf diese Weise konnte systematisch vom Allgemeinen – der Erfahrung von Industriearbeit im Ruhrgebiet – zum unter der Oberfläche der Sammlung schlummernden Besonderen vorgedrungen werden: die Rolle von Frauen, die in der Nachkriegszeit zunehmend in die Verwaltungen der Industrieunternehmen des Ruhrgebiets einzogen, bei der Automatisierung und Computerisierung von Verwaltungsvorgängen. In einem anschließenden *Close Reading* der gefilterten Interviews fielen die nach Geschlecht divergierenden Narrative und Wertungen von Modernisierungsprozessen auf.