

Axel Salheiser

104.1 Was sind Dokumente und was ist Dokumentenanalyse?

104.1.1 Dokumente

Unter *Dokumenten* sollen in diesem Kapitel natürliche Daten verstanden werden, die in schriftlicher Form als Texte vorliegen. Dokumente sind insofern *natürliche Daten*, als dass sie nicht zu Forschungszwecken und ohne die Beteiligung oder Intervention der Forschenden entstanden sind. Für die Sozialwissenschaft sind folgende *Dokumententypen* relevant:

- *Offizielle Dokumente bzw. Publikationen*, z. B. Hand- und Jahrbücher, Bücher (Belletristik, Fachliteratur) (Ernst, Kapitel 105 in diesem Band), Manuskripte öffentlicher Reden, Zeitschriften und Zeitungen (Klein, Kapitel 108 in diesem Band), Broschüren, Flugblätter, Karten usw.;
- *Interne Dokumente aus öffentlichen/staatlichen Verwaltungsorganisationen und privaten Unternehmen/Organisationen* (Meyermann/Gebel/Liebig, Kapitel 103 in diesem Band), z. B. Akten, Statistiken, Listen, d. h. Buchhaltungs- und Registerdaten (Hartmann/Lengerer, Kapitel 106 in diesem Band); Dienst- bzw. Geschäftskorrespondenz, Gesprächsprotokolle/-notizen und andere Arbeitsmaterialien;
- *Egodokumente* (Ernst, Kapitel 105 in diesem Band), d. h. persönliche Unterlagen, wie z. B. Briefe, private Aufzeichnungen und Notizen, Tagebücher, autobiographische Zeugnisse, wie z. B. verschriftlichte Augenzeugenberichte und Gedächtnisprotokolle.

Dokumente können sowohl in gedruckter als auch in handschriftlicher Form auf Papier vorliegen, aber auch auf Mikrofiche, Mikrofilm oder anderen Datenträgern und (elektronischen) Medien gespeichert sein.

Bei Dokumenten handelt es sich um (Träger von) *prozessproduzierten Daten* (Müller 1977: 1): „Unter prozessproduzierten Daten werden all diejenigen Daten verstanden, die als Aufzeichnungen öffentlicher und privater Organisationen im Rahmen ihrer Tätigkeit und nicht nur zum Zwecke wissenschaftlicher bzw. statistischer Auswertung gesammelt werden bzw. wurden.“

Diese Definition wurde in erster Linie für Daten aus der „sozialen Buchführung“ (der öffentlichen Sozialverwaltung, des Personalwesens in Unternehmen usw.; Bick/Müller 1984: 42) gefunden. Sie lässt sich jedoch auch auf schriftliche Quellen ausweiten, die Individualpersonen zuzurechnen sind, z. B. private Aufzeichnungen, Briefe oder andere Ego-Dokumente (Ernst, Kapitel 105 in diesem Band).

Dokumente sind *Artefakte kommunikativer Praxis* (Schubert, Kapitel 113 in diesem Band) und können deshalb eine Datenbasis für sozialwissenschaftliche Interpretationen darstellen (Hitzler/Honer 1997: 8, Prior 2003). Maßgeblich ist, welche Erkenntnisse über soziales Handeln sich aus einem Dokument erschließen lassen und welche sozialen Prozesse und Strukturen daraus rekonstruiert werden können. Beispielsweise ermöglicht die Auswertung von Akten Aussagen über das Verwaltungshandeln, das in ihnen abgebildet ist (Bick/Müller 1984: 74). Dies macht die Dokumentenanalyse u. a. auch für die Praxisforschung in der Sozialen Arbeit (Maschke 2016) und die empirische Erziehungswissenschaft (Glaser 2010) interessant. Eine sozialwissenschaftliche Quellenkritik muss allerdings reflektieren, dass Akten, u. a. amtliche Dokumente, immer auch den Charakter „institutioneller Zurschaustellungen“ (Goffman 1973) haben, in denen sich die Legitimitäts-, Rationalitäts- und Effizienzkalküle ihrer Produzenten und ihrer sozialen Umwelten manifestieren (Wolff 2008).

In der Geschichtswissenschaft spricht man von Dokumenten als *schriftliche Quellen* oder *Zeugnisse*, in denen historische Ereignisse geschildert werden oder historische Sachverhalte zum Ausdruck kommen. Typische Beispiele dafür sind historische Briefe, Chroniken und Urkunden.

104.1.2 Ziele und Varianten der Dokumentenanalyse

Da schriftliche Quellen traditionell die wichtigste Datengrundlage für die Geschichtsforschung sind, kommt einer *Quellenkunde* eine zentrale Bedeutung zu, in der die kritische Bewertung der Authentizität, Verlässlichkeit und Aussagekraft der Quellen gewährleistet wird (Beck/Henning 2003, Rusinek et al. 1992). Bei historischen Dokumenten sind neben der Autorenschaft sowie der zeitlichen und geographischen Einordnung auch der Informationsgehalt und die Bedeutung des Dokumentes zu klären, was üblicherweise im Vergleich mit anderen Quellen geschieht, die als weitestgehend gesichert gelten können.

Doch auch die Erforschung von Gegenwartsgesellschaften und neueren und neuesten sozialen Phänomenen kann sich auf die Auswertung entsprechender schriftlicher Quellen, also Dokumente, stützen. Grundsätzlich unterscheidet sich *sozial-*

wissenschaftliche Dokumentenanalyse von der historischen Quellenarbeit zunächst durch den *zeitlichen Referenzrahmen* und das damit verknüpfte fachspezifische Erkenntnisinteresse. Die Geschichtswissenschaft stellt meist eine narrativ-deskriptive bzw. „philologisch-hermeneutische“ (Weyrauch 1977: 141) Quellenarbeit in den Mittelpunkt, um beispielsweise Ereignis- oder Ideengeschichte zu rekonstruieren. Die Untersuchung historischer Prozesse bzw. von Entwicklungslinien sozialen Wandels (besonders der Neuzeit) hat vor allem innerhalb der historischen Sozialforschung ihre Heimat gefunden, wobei die interdisziplinäre Verschränkung mit der Soziologie auch eine Öffnung der Geschichtswissenschaft für neue Fragestellungen und neue methodische Ansätze bewirkt hat, darunter quantitative bzw. statistische Methoden.

In der modernen Soziologie ist zwar die Befragung (Helfferich, Reinecke, Kapitel 55 und 62 in diesem Band) als „Königsweg“ (Scheuch 1977) der Datenerhebung etabliert, die Dokumentenanalyse (als Sekundärdatenanalyse) kann aber gerade dort wertvolle Impulse liefern, wo Primärerhebungsdaten nicht vorliegen, dem Erkenntnisinteresse entsprechend ergänzt werden sollten oder ein anderweitiger Feldzugang erschwert ist. Abhängig vom Dokumententyp sind sowohl *explorative* als auch *konfirmatorische Dokumentenanalysen* möglich, was auf die Anknüpfungspunkte zu diversen qualitativen (Przyborski/Wohlrab-Sahr, Kapitel 7 in diesem Band) und quantitativen (Stein, Kapitel 8 in diesem Band) Arbeitsweisen und Forschungstraditionen verweist.

Ziel einer sozialwissenschaftlichen Dokumentenanalyse kann es sein, die *Interaktionsmuster, institutionellen Handlungskontexte* sowie „Wertorientierungen oder Meinungen von Individuen oder Gruppen aufzudecken“ (Best 1977: 162). Anschließend an die klassischen Definitionen der *Inhaltsanalyse* soll dabei von manifesten Textinhalten auf latente soziale Kontexte geschlossen werden (Früh 2011, Mayring/Fenzl, Kapitel 43 in diesem Band).

In erster Linie ist entscheidend, in welcher Struktur und in welchem Umfang die Dokumente vorliegen, so dass beispielsweise im Fall der massenhaften Verfügbarkeit des Dokumententyps B (interne Dokumente aus Verwaltungsorganisationen) eine quantitative Analyse (Hartmann/Lengerer, Kapitel 106 in diesem Band) „prozessproduzierter Daten“ (Scheuch 1977) sinnvoll ist. Hingegen dürften bei einzelnen, unter Umständen sogar nur fragmentarischen Dokumenten (vor allem der Dokumententypen A, C) qualitative, d. h. hermeneutische bzw. interpretative (Kurt/Hebrink, Kapitel 38 in diesem Band) und inhaltsanalytische Verfahren (Mayring/Fenzl, Kapitel 43 in diesem Band) angemessen sein. Nachfolgend werden Grundlagen, Anwendungspotentiale und Probleme der Dokumentenanalyse diskutiert.

Prinzipiell ist zu unterscheiden zwischen der primären, *manifesten* Sinnebene des Textes, auf der soziologisch relevante Sachverhalte durch die Darstellungen und Bewertungen des Autors wiedergegeben und vom Forscher direkt „abgelesen“ und interpretiert werden können (z. B. Ablauf und Bewertung einer Interaktionssituation), und sekundären, *latenten* Sinnebenen, auf die aufgrund des Textes geschlossen werden kann (z. B. Bedeutung der Interaktion). Beispielsweise ermöglichen das Vokabu-

lar und der Sprachstil in einem Dokument Aussagen über die Bildung und den Status des Autors, während der geschilderte Sachverhalt eher „unpersönlich“ ist. Strukturierende und zusammenfassende inhaltsanalytische Auswertungsstrategien können hier zwar für viele Fragestellungen zielführend sein, oftmals stellt sich aber die Frage eines tieferen Sinnzusammenhangs zwischen Dokument (Text) und sozialem Umfeld (Kontext). Ebenso wird der Ganzheitlichkeit eines Textes durch Kategorienbildung (Textvercodung) und zusammenfassende bzw. quantifizierende Auswertungen nur selten Rechnung getragen. Ein weiteres Problem konventioneller inhaltsanalytischer Methoden (Mayring/Fenzl, Kapitel 43 in diesem Band) ist, dass eine Bedeutungsäquivalenz der Begriffe im Verständnis von Textproduzent und Analytiker unterstellt wird (Hitzler/Honer 1997: 23). Oftmals lassen sich Sinnmuster in einem Dokument jedoch nur durch eine hermeneutische Textinterpretation, z. B. mittels der Sequenzanalyse, rekonstruieren. Besonders zu beachten sind dabei implizite oder explizite Verweise auf *Außertextliches*, wie metaphorische Anspielungen und charakteristische sprachliche Floskeln. Auch die Bedeutung intendierter oder unintendierter Auslassungen und Verkürzungen kann zentral für ein soziologisches *Textverstehen* (Kurt/Herbrik, Kapitel 38 in diesem Band) sein.

Neben inhaltlichen und sprachlichen Aspekten können außerdem eine Reihe weiterer (formaler) Textgestaltungsmerkmale und Dokumenteneigenschaften sozialwissenschaftlich relevant sein; von der Schriftart und Schriftfarbe über das Trägermaterial, auf dem das Dokument verfasst wurde, bis zum Erhaltungszustand (z. B. Beschädigungen). Die eigentliche *Bedeutung* und der Kontext eines Druckerzeugnisses oder einer maschinengeschriebenen Akte mögen sich erst durch einen eher unscheinbaren handschriftlichen Zusatz oder eine hastige Streichung erschließen. Die Handschrift und Papierqualität eines Briefes wiederum können wichtige Aussagen über den Charakter der Beziehung zwischen Sender und Empfänger sowie allgemeine gesellschaftliche Normen und Konventionen liefern.

104.2 Methodologische Grundlagen der Dokumentenanalyse

Anders als bei der Erhebung von Primärerhebungsdaten aus sozialwissenschaftlichen Befragungen oder in Experimenten (Eifler/Leitgöb, Kapitel 13 in diesem Band) bietet sich mit der Dokumentenanalyse ein *nicht-reaktives Verfahren* (Webb et al. 1975), d. h. die Reaktion der Untersuchungsobjekte wird nicht durch die Datenerhebungstechnik selbst verzerrt. Die außerwissenschaftliche Herkunft der Daten legt allerdings nicht nur dem Historiker, sondern auch dem Sozialwissenschaftler die Pflicht auf, die Validität der Datenbasis zu kontrollieren und zu gewährleisten. Eine Dokumentenanalyse kann niemals nur *datenimmanent* bleiben, sich also allein auf die Dokumenteninhalte berufen, sondern muss mit einer *Rekonstruktion des Entstehungs- und Nutzungskontextes* des Dokumentes einhergehen: „Texte sind als Teile größerer Kommunikationsprozesse aufzufassen. Die Interpretation des Einzeltextes erfordert die

Rekonstruktion der Gesamtstruktur des Kommunikationsprozesses, dessen Teilelement vorliegt“ (Heckmann 1992: 152).

Scheuch (1977: 25) hat auf zwei wichtige Probleme der Nutzung prozessproduzierter Daten für die Forschung verwiesen, nämlich auf

1. die *Selektivität der Daten*, die sich aus dem Entstehungszusammenhang und Verwendungszweck der Daten (z. B. dem Auftrag bzw. der Zuständigkeit einer aktenführenden Verwaltungsbehörde) ergibt, und
2. die *Kategorisierung* bzw. Struktur der *Daten*, die den Erfordernissen der Verwaltungstätigkeit, aber nicht zwangsläufig wissenschaftlich-methodischen Standards entspricht.

Nach dem Verständnis sozialwissenschaftlicher Forschung sind nicht selbst erhobene *Daten* also systematisch verzerrt (Müller 1977: 2, Häder/Häder und Engel/Schmidt, Kapitel 27 und 29 in diesem Band). Um jedoch den Ansprüchen an Reliabilität und Validität (Menold/Krebs, Kapitel 35 in diesem Band) gerecht zu werden, ergibt sich zunächst das Erfordernis einer *Fehlerlehre* (Baur 2009: 21). Diese Fehlerlehre ist analog einer historischen Quellenkritik (Bick/Müller 1984: 42) und muss klären, *zu welchem Zweck* und *wie* die Daten produziert, genutzt und überliefert wurden. Nur so können die *Ursachen* und das *Ausmaß* von Verzerrungen und Datenlücken evaluiert werden, um anschließend Anpassungs- und Korrekturmaßnahmen vorzunehmen.

104.2.1 Umgang mit der Selektivität der Daten

Hinsichtlich der Selektivität der Daten sind besonders jene *systematischen Abbildungsverzerrungen* zu beachten, die sich aus einer heterogenen Datenerfassungs-, -selektions- und -nutzungspraxis innerhalb der datenerzeugenden (aktenführenden) Verwaltungsorganisation ergeben und als „shifting validity“ bezeichnet werden. Insofern entspricht die Fehlerlehre prozessproduzierter Daten einer „organisationssoziologischen Quellenkritik“ (Heckmann 1992: 154):

1. Einerseits können *innerhalb einer Organisation* ganz *unterschiedliche Dokumentations- und Aufbewahrungsvorschriften* bzw. *Zuständigkeiten der Informationsverarbeitung* existieren, denen zudem im betrieblichen bzw. Behörden-Alltag unterschiedlich folgegeleistet wird. Nicht alle Verwaltungsvorgänge sind dokumentationspflichtig oder werden als dokumentationswürdig betrachtet und werden deshalb (in gleichem Umfang) aktenkundig. Aber auch Manipulationen von Buchhaltungsunterlagen bzw. Berichtsfälschungen (vor allem statistischer Angaben) können relevant sein (Webb et al. 1975: 79). Unter Umständen existiert eine administrative Geheimhaltungspraxis, die zu interner Intransparenz und einer unvollständigen Überlieferung von Dokumenten bzw. Daten führt. Beispielswei-

se kann bei Personalakten die Verarbeitung und Kontrolle den Mitarbeitern eine unterschiedliche Priorität eingeräumt werden. So kommt es schließlich dazu, dass die Vollständigkeit, die Aussagekraft und die Fehleranfälligkeit von personenbezogenen Angaben über Funktionsträger von ihrer vertikalen und horizontalen Position in der Organisationshierarchie abhängen (Salheiser 2009).

2. Andererseits sind *technische Probleme bei der Datenerfassung* (z. B. die Störanfälligkeit von EDV-Anlagen), „subjektive“ *Opportunitätsstrukturen und die individuelle Motivation des Personals* für die Datenqualität relevant (z. B. Alltagsroutinen der Eingabe-, Kodier- und Kontrollpraxis, mangelnde Sorgfalt oder Kompetenz usw.), sie sind mitunter eine unsystematische und nur schwer zu rekonstruierende Fehlerquelle.

Um die Fehleranfälligkeit und Phänomene der „shifting validity“ beurteilen zu können, sind *Strategien der externen Validierung* unabdingbar: So muss neben einer *Analyse von Vergleichsdaten* (sofern verfügbar) auch *Kontextwissen* über den primären Entstehungs- und Nutzungsprozess der Daten und die Struktur der aktenführenden Verwaltungsorganisation generiert und genutzt werden. Dies erfordert i. d. R. die Analyse weiterer Dokumente (Akten, Organigramme usw.) sowie die externe Validierung anhand von Vergleichsstatistiken. Gegebenenfalls können Experteninterviews (Helfferich, Kapitel 55 in diesem Band) mit solchen Personen hilfreich sein, die selbst an der Generierung der Daten bzw. der Erstellung der Dokumente beteiligt waren.

104.2.2 Umgang mit der Struktur (Kategorisierung) der Daten

Bezüglich der Kategorisierung prozessproduzierter Daten ist wichtig, dass in vielen Fällen keine unmittelbare Kompatibilität mit den Standards sozialwissenschaftlicher Messung (Stein, Krebs/Menold und Faulbaum, Kapitel 8, 35 und 38 in diesem Band) gegeben ist. Die in den Dokumenten verwendeten *Begriffe* können beispielsweise unpräzise oder widersprüchlich sein und einem spezifischen Verwaltungsjargon unterliegen. Oftmals folgen *Merkmale und Kategorienschemata* (z. B. zur Erfassung personenbezogener Daten) der Eigenlogik der Verwaltungsorganisation. Die Aussagekraft der Daten für wissenschaftliche Auswertungen kann beeinträchtigt sein, wenn bei der Datenerfassung unterschiedliche Merkmale oder Kategorien zusammengefasst wurden (Informationsverlust). Andererseits können bestimmte Merkmale, Merkmalsausprägungen bzw. Unterkategorien existieren, die für Forschungszwecke irrelevant sind (redundante Informationen). Möglicherweise sind zusammengehörende Angaben auch über verschiedene Teildatensätze verstreut, so dass sie aus unterschiedlichen Merkmalen extrahiert, neu geordnet und zusammengeführt werden müssen, um vollständige und konsistente Informationen zu gewinnen.

Die Verwendung von Begriffen und Kategorienschemata kann außerdem nach politisch-ideologischen oder strategischen Vorgaben erfolgt sein (u. a. zur Verschlei-

rung unerwünschter Sachverhalte etc.; Salheiser 2009), was eine direkte Übernahme der Daten bzw. der Datenstruktur für wissenschaftliche Auswertungszwecke und eine neutrale Beurteilung unmöglich macht.

Die Sprache der Dokumente ist per se nicht die Sprache der Wissenschaft, sondern sie unterliegt den Konventionen und der Logik des sozialen Kontextes, aus dem die Dokumente stammen, und bedarf deshalb einer Übersetzung. Daraus ergibt sich die praktische Notwendigkeit einer nachträglichen Operationalisierung (Stein, Kapitel 8 in diesem Band), was der Analyse prozessproduzierter Daten eine Sonderstellung innerhalb sozialwissenschaftlicher Ex-post-facto-Forschungsdesigns (Stein, Kapitel 8 in diesem Band) einräumt. Dabei gilt es, die aufgefundenen Merkmalskategorien der sozialen Buchführung in brauchbare Indikatoren zu überführen und die prozessproduzierten Daten dementsprechend datentechnisch aufzubereiten und umzustrukturieren.

Unter Umständen treten bei historischen prozessproduzierten Daten Konsistenzprobleme auf, wenn Begriffe, Bezeichnungen, Maßeinheiten, statistische Kennziffern usw. während der Datenproduktion über längere Zeiträume einem Bedeutungswandel unterlagen und Kodierschemata aus technischen oder organisatorischen Gründen abgeändert wurden. Hier ist prozessbezogenes Kontextwissen unerlässlich, um die Reliabilität und Validität der Daten zu kontrollieren und sicherzustellen.

Auch Dokumente, deren Inhalte als „qualitative Daten“ zu klassifizieren sind, bedürfen zunächst einer grundlegenden Validierung und Plausibilisierung (Flick, Kapitel 34 in diesem Band). Betrachtet man beispielsweise einen persönlichen Bericht als Zeugnis sozialen Handelns, so sind, wie in der historischen Quellenkunde, der zeitliche, der geographische und der soziale Entstehungskontext des Berichtes sowie die Autorität, Kompetenz und Glaubwürdigkeit des Verfassers einzuordnen bzw. zu beurteilen (Howell/Prevenier 2004: 82 ff.). Wo lassen sich vergleichbare Angaben nachweisen, wo treten Widersprüche zu tage? Inwieweit korrespondieren die Dokumente mit anderen Quellen, gibt es etwa spiegelbildlich überlieferte Bestände, die sich aus internen Querverweisen und Belegen rekonstruieren lassen (z. B. Verteiler-Abkürzungen auf Verwaltungsakten)?

Briefwechsel, die nur auf der Empfängerseite überliefert sind, gestatten mitunter erst durch die Zusammenführung aus verschiedenen Beständen und durch eine komparative Textanalyse valide Aussagen über den Forschungsgegenstand. Grundsätzlich gilt: Keine Quelle und kein Dokument kann Anspruch auf Objektivität erheben (Menold/Krebs, Flick, Herbrink/Kurt, Kapitel 35, 34 und 38 in diesem Band). Jede Wiedergabe von Sachverhalten, mag sie auch noch so neutral anmuten, unterliegt einem Konstruktionsprozess der Wahrnehmung und der kognitiven Verarbeitung und bildet einen der Intention, Motivation und Emotion des Autors entsprechend gefilterten *Ausschnitt* sozialer Realität ab, sie ist also subjektiv gefärbt und „verzerrt“.

104.3 Datenzugang

Wie alle anderen Datentypen können auch Dokumente nur einen Ausschnitt (subjektiv wahrgenommener) sozialer Wirklichkeit wiedergeben. Forschungspraktisch ergibt sich zunächst die Problematik des Zugangs zu geeigneten Dokumenten, insofern sie nicht ohne weiteres als Publikationen (Dokumententyp 1) verfügbar sind.

104.3.1 Erhebung zusammen mit anderen Daten

Persönliche Dokumente (Dokumententyp 3), die in mikrosoziologischen Forschungskontexten relevant sind, wie z. B. Egodokumente, Briefe, Urkunden, werden in der Regel im Zusammenhang mit dem Feldzugang anderer Datenerhebungsmethoden – der narrativen Interviewführung (Küsters, Kapitel 56 in diesem Band), der teilnehmenden Beobachtung (Thierbach/Petschick, Knoblauch/Vollmer, Kapitel 109 und 41 in diesem Band) usw. – verfügbar, zumindest wenn sie im sozialen Umfeld des Verfassers und seiner Interaktionspartner verblieben sind. Anderenfalls stellt sich die Frage nach der Archivwürdigkeit bzw. einer archivischen Überlieferung.

104.3.2 Dokumente im Privatbesitz

Ein persönliches Dokument kann aufgrund seines Inhaltes (z. B. des biographischen und familienbiographischen Bezugs) von hohem ideellem Wert für seinen Besitzer sein. Werden die Einsicht, die Reproduktion, die vorübergehende oder gar dauerhafte Überlassung bzw. der Erwerb solcher Dokumente beabsichtigt, sollte der Forscher besonderes Augenmerk auf die Vermittlung von Vertrauenswürdigkeit und der Seriosität seines Vorhabens legen. Eine schriftliche Übereinkunft über die Nutzung der Dokumente (insbesondere von Angehörigen bzw. von Hinterbliebenen) sollte auf jeden Fall vorliegen, auch wenn die Veröffentlichung der Dokumente bzw. der Forschungsergebnisse in der Regel als völlig unproblematisch eingeschätzt wird. Erfahrungsgemäß kann sich jedoch während des Forschungsprozesses die Kooperationsbereitschaft derer ändern, die persönliche Dokumente von sich oder aus dem Nachlass verstorbener Angehöriger zur Verfügung gestellt haben. Dem ausdrücklichen Wunsch der Nichtveröffentlichung zuwiderzuhandeln, könnte einen ernstzunehmenden Verstoß gegen Prinzipien der Forschungsethik darstellen und u. U. sogar straf- und zivilrechtliche Konsequenzen nach sich ziehen.

104.3.3 Dokumente aus Organisationen

Die Verfügbarkeit von internen Dokumenten öffentlicher und privater Organisationen (Dokumententyp 2) für eine wissenschaftliche Nachnutzung, darunter auch der Zugriff auf prozessproduzierte Massendaten, ist zunächst abhängig von der Kooperationsbereitschaft der Organisationen (Meyermann/Giebel/Liebig, Kapitel 103 in diesem Band) und von übergeordneten datenschutzrechtlichen Belangen (Mühlichen, Kapitel 23 in diesem Band). Insbesondere bei historischen Dokumenten stellt sich aber auch die Frage nach einer archivischen Überlieferung, der Vollständigkeit und dem Erhaltungszustand solcher Daten (Webb et al. 1975: 80). Wie Historiker wissen, kann die Erschließung archivalischer Datenquellen deshalb Detektivarbeit sein. Vor allem der Einsicht in umfangreiche Aktenbestände bzw. von prozessproduzierten Massendaten sollte eine gründliche Planung vorausgehen.

104.3.4 Dokumente aus Archiven und Dokumentationsstellen

Öffentliche und private Archive oder Dokumentationsstellen gehören zu den wichtigsten Institutionen, die Dokumente zur Verfügung stellen. Das geeignete Archiv muss entsprechend seiner sachlichen und räumlichen Zuständigkeit ausgewählt werden. Auf nationaler und internationaler Ebene stehen dafür Archivverzeichnisse und Online-Portale zur Verfügung (Arathymou 2009). Archive sind in der Regel an die Verwaltungen bzw. Einrichtungen angegliedert, deren Akten sie aufbewahren. So ist zwischen Archiven des Bundes, den Landesarchiven, kommunalen Archiven, Archiven von Universitäten und anderen wissenschaftlichen Einrichtungen sowie Archiven von Wirtschaftsunternehmen, Verbänden, Stiftungen und Vereinen zu unterscheiden. Bei historischen Akten muss beachtet werden, dass die Zuständigkeit des Archivs von einer Rechtsnachfolge der jeweiligen Organisation bzw. politischen Einheit (Staat, Bundesland, Kommune) und der Territorialgeschichte abhängen kann.

Der Zugang zu staatlichen Archiven in Deutschland ist in den Archivgesetzen des Bundes und der Länder geregelt. Bei privaten Archiven liegt es hingegen im Ermessensspielraum der Organisationen, inwiefern sie ihre Archivalien zur wissenschaftlichen Nutzung bereitstellen. Öffentliche und private Archive können sich außerdem erheblich bezüglich der Nutzungsbedingungen unterscheiden. Dies betrifft insbesondere die Veröffentlichung der gewonnenen Befunde bzw. die Reproduktion der benutzten Dokumente. So muss beispielsweise ein wissenschaftliches Interesse an größtmöglicher Transparenz mit den Interessen eines Unternehmens an seinem Image in der Öffentlichkeit abgestimmt werden. Es ist unabdinglich, sich im Vorfeld eines Forschungsvorhabens möglichst umfangreich über die allgemeinen Archivnutzungsvorschriften, die Bedingungen der Akteneinsicht und die verfügbaren *Findmittel* zu informieren. (*Findmittel* sind schriftliche oder elektronische Verzeichnisse des Aktenbestandes.)

Staatliche bzw. öffentliche Archive wie das Bundesarchiv stellen auf ihren Internetseiten hilfreiche Angaben zur *Archivstruktur*, *Bestandsübersichten* und weitreichende elektronische Findmittel zur Verfügung. Bei kleineren öffentlichen Archiven und Unternehmensarchiven erfolgt die Auskunft üblicherweise mittels schriftlicher Kontaktaufnahme mit den Archivmitarbeitern bzw. einer wissenschaftlichen Begleitinstanz der archivführenden Institution.

Die Einsicht und die Nutzung von Archivgut bedürfen in der Regel einer vorherigen *Anmeldung bzw. Nutzungserlaubnis*. Dabei sind mitunter Sonderregelungen und Nutzungseinschränkungen für Teilbestände zu beachten, die (a) sensible Daten enthalten und damit besonderen Datenschutzvorschriften (z. B. Sperrfristen) unterliegen oder (b) aus archivischen Gesichtspunkten (Alter, Erhaltungszustand, historische Bedeutung) als besonders schützenswert beurteilt werden.

In vielen Fällen ist ein zeitlicher Abstand zwischen Anmeldung und gewünschtem Zeitpunkt der Dokumenteneinsicht von mehreren Tagen oder sogar Wochen notwendig, da das Archivgut zur Nutzung bereitgestellt (ausgehoben) und ggfs. zuvor von den Archiv-Mitarbeitern gesichtet und aufbereitet werden muss, bevor es dem Nutzer vorgelegt werden kann. Im erforderlichen *Nutzungsantrag* sollte das Forschungsvorhaben knapp geschildert werden, wobei insbesondere darzustellen ist, in welcher Form eine Veröffentlichung der Forschungsergebnisse beabsichtigt ist und inwieweit später die Reproduktion des Archivgutes (z. B. die Verwendung von Fotos aus dem Archiv in der geplanten Publikation) gewünscht ist.

Viele Archive machen hier Einschränkungen bzw. Auflagen gemäß ihren Bild- bzw. Urheberrechten und dem Datenschutz, insbesondere wenn es sich um sensible, personenbezogene Unterlagen (z. B. Personalakten von Verwaltungsorganisationen oder private Korrespondenz lebender Personen) handelt. Gerade im Fall biographischer Analysen (Rosenthal, Kapitel 40 in diesem Band) oder der Rekonstruktion einzelfallbezogenen Verwaltungshandelns (Hering/Jungmann, Kapitel 42 in diesem Band) kann eine Anonymisierung dem wissenschaftlichen Erkenntnisinteresse und dem methodischen Zugang zuwiderlaufen (Wolff 1995). Zu Zwecken wissenschaftlicher Forschung können deshalb Ausnahmeregelungen getroffen werden, solange den Anforderungen der allgemeinen Datenschutzaufgaben entsprochen wird. In der Regel ist vor der Veröffentlichung eine Anonymisierung bzw. Pseudonymisierung personenbezogener Unterlagen zu gewährleisten. Insofern die Anonymisierung nicht bereits von der aktenausgebenden Stelle vorgenommen wurde, hat der Forscher dafür Sorge zu leisten, dass die Persönlichkeitsrechte der in den Dokumenten erwähnten Personen gewahrt bleiben. Dies kann einerseits durch eine gründliche Unkenntlichmachung von Namen erfolgen, andererseits kann es erforderlich sein, im Text auf die Schilderung charakteristischer privater Details zu verzichten, die eine Identifikation der entsprechenden Personen durch den Leser ermöglichen würde. Im Zweifelsfall ist vor der Veröffentlichung eine Rücksprache mit den Archivmitarbeitern bzw. eine Beratung mit anderen wissenschaftlichen Experten geboten.

Eine wichtige Ausnahme von der grundsätzlichen Anonymisierungspflicht bei der Veröffentlichung personenbezogener Unterlagen ergibt sich in Bezug auf sogenannte „*Personen der Zeitgeschichte*“. Darunter werden Personen verstanden, die aufgrund ihrer herausragenden gesellschaftlichen Bedeutung im Fokus öffentlichen Interesses und medialer Aufmerksamkeit stehen, wie z. B. Politiker oder prominente Künstler. Erstens zeigt diese Definition jedoch, dass es interpretationsbedürftige Grenzfälle geben kann, und zweitens sind in jedem Fall der Schutz der Privat- und Intimsphäre der betreffenden Personen zu wahren. Im Sinne der wissenschaftlichen Forschung und der Information der Öffentlichkeit ergibt sich prinzipiell jedoch ein ausreichender Spielraum für die wissenschaftliche Verwertung personenbezogener Dokumente, wenn die darin dargestellten Sachverhalte von öffentlichem Belang sind.

Viele Archive bieten ihren Nutzern die grundsätzliche Möglichkeit, sich im Anschluss an die Akteneinsicht *Fotokopien von Auszügen des Archivgutes* erstellen und diese zuschicken zu lassen. Dies hat den Vorteil, dass eine intensive Dokumentenanalyse nach einer ersten Sichtung des Materials auch unabhängig vom Archivstandort erfolgen kann. In jedem Fall muss in einer Publikation, die sich auf Archivgut stützt, ein sorgfältiger und vollständiger Nachweis der verwendeten Quellen erfolgen (Angabe des Archivs sowie der Signaturen, d. h. der archivinternen Bestandskürzel und Bestellnummern zur Dokumentenidentifikation).

104.4 Auswahl von Dokumenten

Während man üblicherweise kleine Dokumentenbestände insofern hypothesen- und materialgeleitet sichtet, dass man während der Einsicht die Auswahl weiterer Teilbestände und Einzeldokumente in Abhängigkeit von deren Relevanz für die Forschungsfrage trifft (Akremi, Kapitel 26 in diesem Band), kann bei größeren Beständen, z. B. bei alphabetischen Personenregistern, eine systematische Zufallsauswahl (random sampling) der Dokumente notwendig sein. Dabei können Verfahren der Stichprobenziehung zum Einsatz kommen, die denen der schriftlichen oder mündlichen Befragung (Häder/Häder, Kapitel 27 in diesem Band) ähneln. So könnte beispielsweise eine bestimmte Anzahl von Akten festgelegt werden, die jedem n-ten Aktenbehälter (Karteikasten o. ä.) zu entnehmen ist, bis eine ausreichende Stichprobengröße erreicht worden ist (vgl. Buchholz 2002).

104.5 Analysestrategien und Anwendungsmöglichkeiten

104.5.1 Methoden-Mix

Bei der sozialwissenschaftlichen Dokumentenanalyse gilt es, je nach der Beschaffenheit der Dokumente und dem jeweiligen Erkenntnisinteresse entsprechend, etablierte Verfahren der qualitativen und quantitativen Sozialforschung und anderen Sozial- und Geisteswissenschaften zu adaptieren. Sowohl die ausschließliche Analyse von Dokumenten als auch die Verschränkung der Dokumentenanalyse mit anderen Forschungsstrategien (Methoden-Mix, Kelle, Kapitel 9 in diesem Band) sind möglich.

104.5.2 Qualitative Inhaltsanalyse

So kann eine qualitative Inhaltsanalyse (Mayring/Fenzl, Kapitel 43 in diesem Band) von schriftlichen Zeugnissen aus dem zu untersuchenden sozialen Gegenstandsreich als unterstützendes, ergänzendes Verfahren im Zusammenspiel mit der qualitativen oder quantitativen Analyse von Befragungsdaten (Reinecke und Helfferich, Kapitel 62 und 55 in diesem Band) oder Beobachtungsdaten (Thierbach/Petschick, Kapitel 109 in diesem Band) angewandt werden, um sich in der Phase der Exploration und Hypothesengenerierung den Forschungsgegenstand näher zu erschließen oder um Fallbeispiele (Hering/Jungmann, Kapitel 42 in diesem Band) zu extrahieren, die die Grundlage für weitere Forschungstätigkeit bilden. Hier wäre auch eine Kombination ganz unterschiedlicher Dokumenten- bzw. Texttypen denkbar. Beispielsweise könnten für stadtsoziologische Untersuchungen (Dangschat/Kogler, Kapitel 115 in diesem Band) zur räumlichen Konzentration bzw. Segregation sozialer Gruppen im Vorfeld oder als Ergänzung zu einer repräsentativen Befragung auch öffentliche Broschüren der Stadt, Verwaltungsstatistiken zur Soziodemographie der Bewohner (Auszüge aus dem Einwohnermelderegister), Telefon- und Adressbücher, Kartenmaterial des Katasteramtes zur städtischen Siedlungsstruktur und Raumplanung (Lakes, Kapitel 118 in diesem Band) sowie die Regionalpresse und Stadtteilzeitungen gesichtet und quantitativ oder qualitativ ausgewertet werden.

Ein anderes Beispiel für die Dokumentenanalyse von Publikationen bietet eine qualitative Inhaltsanalyse von Populärliteratur mit dem Ziel, den Wandel gesellschaftlicher Diskurse und kultureller Normen (u. a. von Geschlechterverhältnissen und Rollenvorstellungen) zu untersuchen (Ernst 2009).

104.5.3 Biographieforschung und sozialwissenschaftliche Hermeneutik

Für mikrosoziologische Einzelfallstudien (Hering/Jungmann, Kapitel 42 in diesem Band), z. B. der Familien- oder Biographieforschung (Rosenthal, Kapitel 40 in diesem Band), ist denkbar, in einem mehrstufigen Prozess (im Sinne eines „theoretical samplings“ im Rahmen der Grounded Theory; Strübing, Kapitel 37 in diesem Band) (Ego-)Dokumente (Ernst, Kapitel 105 in diesem Band) und Transkripte narrativer Interviews (Küsters, Kapitel 56 in diesem Band) miteinander zu vergleichen und dabei bereits gewonnene Befunde zu überprüfen und weiterzuentwickeln. Dem Forschungsgegenstand und dem korrespondierenden Datenmaterial gemäß ergibt sich ein *interpretativer* und *verstehender* Zugang zu den Dokumenten (Kurt/Herbrik, Kapitel 38 in diesem Band). Die „Implikationen unterschiedlicher Gestaltungsformen und Darstellungsstrategien“ (Wolff 2008) von Dokumenten können hier besonders relevant sein. Bei einer biographischen Fallrekonstruktion können Ego-Dokumente zur Familiengeschichte (Briefe, Tagebücher) helfen, das in narrativen Interviews Geschilderte einzuordnen, oder aus ihnen lassen sich wertvolle Hinweise für weitere Interviews extrahieren. Die Sequenzanalyse (Kurt/Herbrik, Kapitel 38 in diesem Band) eines Briefes oder Tagebuchauszugs etwa erfolgt dabei ähnlich der eines Interviews.

104.5.4 Quantitative Analysestrategien

Quantitative Analysestrategien für Dokumente, die in der methodischen Tradition indikatorengestützter empirischer Sozialforschung (Burzan und Huinink, Kapitel 47 und 48 in diesem Band) stehen, bieten sich im Fall „von in großen Mengen vorliegenden Quellengruppen wie zum Beispiel Personalakten, Steuerlisten und Kirchenbücher“ (Best/Mann 1977: 1) an.

Bei maschinengeschriebenen Akten (Hartmann/Lengerer, Kapitel 106 in diesem Band) sowie Büchern, Zeitungen (Klein, Kapitel 108 in diesem Band) und anderen gedruckten Dokumenten kann beispielsweise eine *computerunterstützte Inhaltsanalyse* (CUI, Rössler/Geise 2013) stattfinden, insofern sich das Datenmaterial digitalisieren, d. h. scannen und mit automatischen Texterkennungsprogrammen auslesen lässt (Kuckartz/Rädiker, Kapitel 32 in diesem Band).

Enthalten Dokumente eine Vielzahl von (numerischen) Informationen, die homogenen Erfassungsmustern folgen, also Daten in Tabellen, Listen oder Formularen, eignen sich diese besonders für quantitative Analysen. So erlauben etwa die Geburts-, Tauf- und Sterbelisten der Kirchgemeinden mit ihren personenbezogenen Angaben zu Lebensdaten, sozialer und geographischer Herkunft, dem Beruf, dem gesellschaftlichem Status usw. die Rekonstruktion lokaler Sozialstrukturentwicklung eine *Verlaufsdatenanalyse* (Pötter/Prein, Kapitel 52 in diesem Band), sind doch die Veränderungen der Mortalitäts- und Geburtenraten und des Heiratsalters aussagekräftige Indikatoren langfristigen sozialen bzw. kulturellen Wandels (Best/Mann 1977).

Historische Personalakten (die u. a. auch Ego-Dokumente wie hand- oder maschinenschriftliche Lebensläufe enthalten) können insbesondere bei erheblichem Arbeitsaufwand der Aktenrecherche, Kategorienbildung, Vercodung (Züll/Menold, Kapitel 75 in diesem Band), (manuellen) Datenerfassung und -aufbereitung zu statistischen Datensätzen aggregiert werden, die dann beispielsweise die multivariate Analyse von Kollektivbiographien in der empirischen Elitenforschung ermöglichen (Best/Hornbostel 1998). Die *Aggregation von biographischen Individualdaten* aus Dokumenten (Graeff, Kapitel 102 in diesem Band) bedarf einer intensiven Informationsselektion und -reduktion. Die Datenerhebung kann besonders dann sehr aufwändig sein, wenn Daten zu Einzelpersonen aus verschiedenen Quellen fusioniert, ergänzt und abgeglichen werden müssen (ebd.) und wenn es sich teilweise um handschriftliche Dokumente handelt (Best 1977).

Literatur

- Arathymou, Spyridoula (2009): Finding and Accessing the Right Archive and Archival Data. In: HSR 34 (3): 71–77
- Baur, Nina (2009): Measurement and Selection Bias in Longitudinal Data. In: HSR 34 (3): 9–50
- Beck, Friedrich/Henning, Eckart (Hg.) (2003): Die archivalischen Quellen. Köln et al.: Böhlau
- Best, Heinrich (1977): Die quantitative Analyse inhaltlicher und kontextueller Merkmale historischer Dokumente. In: Best, Heinrich/Mann, Reinhard (Hg.): Quantitative Methoden in der historisch-sozialwissenschaftlichen Forschung. Stuttgart: Klett-Cotta, 162–205
- Best, Heinrich/Mann, Reinhard (Hg.) (1977): Quantitative Methoden in der historisch-sozialwissenschaftlichen Forschung. Stuttgart: Klett-Cotta
- Best, Heinrich/Hornbostel, Stefan (1998): Prozeß-produzierte Daten als empirisches Material für eine Soziologie des realen Sozialismus. In: GESIS (Hg.): Materialien zur Erforschung der DDR-Gesellschaft. Opladen: Leske + Budrich, 201–221
- Bick, Wolfgang/Paul J. Müller (1984): Sozialwissenschaftliche Datenkunde für prozeß-produzierte Daten. In: Müller, Paul J./Mann, Reinhard (Hg.): Sozialforschung und Verwaltungsdaten. Stuttgart: Klett-Cotta. 123–159
- Buchholz, Matthias (2002): Stichprobenverfahren bei massenhaft gleichförmigen Einzelfallakten. In: HSR 27 (2/3): 100–223
- Ernst, Stefanie (2009): Using Qualitative Content Analysis of Popular Literature for Uncovering Long-Term Social Processes. In: HSR 34 (1): 252–269
- Früh, Werner (2011): Inhaltsanalyse: Theorie und Praxis. 7., überarb. Auflage. Konstanz et al.: UVK-Verlags-Gesellschaft
- Glaser, Edith (2010): Dokumentenanalyse und Quellenkritik. In: Friebertshäuser, Barbara/Langer, Antje/Prengel, Annedore (Hg.): Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft. 3., vollst. überarb. Auflage. Weinheim: Juventa, 365–375
- Goffman, Erving (1973): Asyle. Über die soziale Situation psychiatrischer Patienten und anderer Insassen. Frankfurt a. M.: Suhrkamp
- Heckmann, Friedrich (1992): Interpretationsregeln zur Auswertung qualitativer Interviews und sozialwissenschaftlich relevanter „Texte“. In: Hoffmeyer-Zlotnik, Jürgen H. P. (Hg.): Analyse verbaler Daten. Opladen: Westdeutscher Verlag, 142–167
- Hitzler, Ronald/Honer, Anne (Hg.) (1997): Sozialwissenschaftliche Hermeneutik. Opladen: Leske + Budrich UTB
- Howell, Martha/Prevenier, Walter (2004): Werkstatt des Historikers. Köln/Weimar/Wien: Böhlau UTB

- Maschke, Dietmar (2016): Dokumentenanalyse. In: König, Joachim (Hg.): Praxisforschung in der Sozialen Arbeit: Ein Lehr- und Arbeitsbuch. Stuttgart: Kohlhammer, 144–148
- Müller, Paul J. (Hg.) (1977): Die Analyse prozeß-produzierter Daten. Stuttgart: Klett-Cotta
- Prior, Lindsay (2003): Using Documents in Social Research. London: Sage
- Rössler, Patrick/Geise, Stephanie (2013): Standardisierte Inhaltsanalyse. In: Möhring, Wiebke/Schlütz, Daniela (Hg.): Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft. Wiesbaden: VS Verlag, 269–287
- Rusinek, Bernd-A./Ackermann, Volker/Engelbrecht, Jörg (Hg.) (1992): Einführung in die Interpretation historischer Quellen: Schwerpunkt: Neuzeit. Paderborn et al.: Schöningh
- Salheiser, Axel (2009): Handling Ideological Bias and Shifting Validity of Longitudinal Data. In: HSR 34 (1): 197–210
- Scheuch, Erwin K. (1977): Die wechselnde Datenbasis der Soziologie – Zur Interaktion zwischen Theorie und Empirie. In: Müller, Paul J. (Hg.): Die Analyse prozeß-produzierter Daten. Stuttgart: Klett-Cotta. 5–41
- Weyrauch, Erdmann (1977): Datenverarbeitung als Quellenkritik? In: Müller, Paul J. (Hg.): Die Analyse prozeß-produzierter Daten. Stuttgart: Klett-Cotta, 141–197
- Webb, Eugene J./Campbell, Donald T./Schwartz, Richard D./Sechrest, Lee (1975): Nichtreaktive Messverfahren. Weinheim und Basel: Beltz
- Wolff, Stephan (1995): Text und Schuld. Die Rhetorik psychiatrischer Gerichtsgutachten. Berlin: de Gruyter
- Wolff, Stephan (2008): Dokumenten- und Aktenanalyse. In: Flick, Uwe/von Kardorff, Ernst/Steinke, Ines (Hg.): Qualitative Forschung: Ein Handbuch. 6. durchgesehene und aktualisierte Auflage. Reinbek: Rowohlt, 502–514

Axel Salheiser ist wissenschaftlicher Referent am Institut für Demokratie und Zivilgesellschaft (IDZ) Jena und Mitglied des KomRex – Zentrum für Rechts-extremismusforschung, Demokratiebildung und gesellschaftliche Integration der Friedrich-Schiller-Universität Jena. *Ausgewählte Publikationen:* Parteitreu, plangemäß, professionell? Rekrutierungsmuster und Karriereverläufe von DDR-Industriekadern. Wiesbaden: VS Verlag (2009); Elites and Social Change – The Socialist and Post-Socialist Experience. Hamburg: Kraemer Publishers (zusammen mit Heinrich Best und Ronald Gebauer, 2009); Handling Ideological Bias and Shifting Validity of Longitudinal Data: The Case of Process-Generated Data on GDR Elites, in: Nina Baur (Hg.): Linking Theory and Data. Process-Generated and Longitudinal Data for Analyzing Long-Term Social Processes. Historical Social Research 34, 1 (2009). *Webseite:* <http://www.idz-jena.de/>. *Kontakt-adresse:* axel.salheiser@idz-jena.de.