

HUMBOLDT-UNIVERSITÄT ZU BERLIN
INSTITUT FÜR BIBLIOTHEKS- UND INFORMATIONSWISSENSCHAFT



BERLINER HANDREICHUNGEN
ZUR BIBLIOTHEKS- UND
INFORMATIONSWISSENSCHAFT

HEFT 475

SPEZIFIKA UND HERAUSFORDERUNGEN QUALITATIVER
DATEN IM FORSCHUNGSDATENMANAGEMENT

VON
MORITZ STRICKERT

SPEZIFIKA UND HERAUSFORDERUNGEN QUALITATIVER
DATEN IM FORSCHUNGSDATENMANAGEMENT

VON
MORITZ STRICKERT

Berliner Handreichungen zur
Bibliotheks- und Informationswissenschaft

Begründet von Peter Zahn
Herausgegeben von
Vivien Petras
Humboldt-Universität zu Berlin

Heft 475

Strickert, Moritz

Spezifika und Herausforderungen qualitativer Daten im
Forschungsdatenmanagement / von Moritz Strickert. - Berlin : Institut für
Bibliotheks- und Informationswissenschaft der Humboldt-Universität zu
Berlin, 2021. S. 122: graph. Darst. - (Berliner Handreichungen zur
Bibliotheks- und Informationswissenschaft ; 475)

ISSN 14 38-76 62

Abstract:

Der Austausch und die (Nach-)Nutzung von Forschungsdaten sind seit Anfang der 2000er-Jahre verstärkt ein wissenschafts(-politisches) Thema. Beschäftigten sich zunächst eher die Fächer, die viele (quantitative) Messdaten produzieren, damit, wie mit den von ihnen geschaffenen und verarbeiteten Daten verfahren werden soll, sind diese Fragen mittlerweile flächendeckend in fast allen Wissenschaftsdisziplinen von Relevanz. Die vorliegende Studie befasst sich mit qualitativen sozialwissenschaftlichen Forschungsdaten und basiert auf leitfadengestützten Expert*inneninterviews, die mit Datenarchiven geführt wurden. Es wird dargestellt, welche Probleme sich auf (forschungs-)ethischer, praktischer und rechtlicher Ebene im Zuge der Archivierung und Nachnutzung dieser Daten ergeben und inwiefern die befragten Einrichtungen mit diesen Problemen umgehen.

Diese Veröffentlichung geht zurück auf eine Masterarbeit im weiterbildenden Masterstudiengang im Fernstudium Bibliotheks- und Informationswissenschaft (Library and Information Science, M. A. (LIS)) an der Humboldt-Universität zu Berlin.

Eine Online-Version ist auf dem edoc Publikationsserver der Humboldt-Universität zu Berlin verfügbar.



Dieses Werk ist lizenziert unter einer [Creative Commons Namensnennung - Nicht kommerziell - Keine Bearbeitungen 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/) Lizenz

Inhalt

1.) EINLEITUNG	8
1.1) PROBLEMATISIERUNG UND FRAGESTELLUNG	8
1.2) AUFBAU DER ARBEIT	12
2.) FORSCHUNGSDATEN(-MANAGEMENT): DEFINITION UND WISSENSCHAFTSPOLITISCHE POSITIONIERUNGEN	13
2.1) DEFINITION VON FORSCHUNGSDATEN	13
2.2) ALLGEMEINE WISSENSCHAFTSPOLITISCHE POSITIONIERUNGEN	15
2.3) SPEZIFIK QUALITATIVER SOZIALWISSENSCHAFTLICHER DATEN UND DEREN NACHNUTZUNG	20
2.4) STELLUNGNAHMEN DER FACHGESELLSCHAFTEN	24
2.5) ETABLIERUNGSPROBLEME DES FORSCHUNGSDATENMANAGEMENTS UND KRITIK AM FORSCHUNGS-DATENPARADIGMA	27
3.) METHODISCHES VORGEHEN	32
3.1) MATERIAL UND METHODIK (FORSCHUNGSDESIGN UND OPERATIONALISIERUNG DER FRAGESTELLUNG)	32
3.1.1) <i>Das Expert_inneninterview: Expert_innenwissen und Expert_innenauswahl</i>	33
3.2) VORGEHENSWEISE EXPERT_INNENINTERVIEWS	34
3.2.1) <i>Feldzugang und Interviewdurchführung</i>	34
3.2.2) <i>Erhebungsinstrument Leitfaden</i>	35
3.3) AUSWERTUNG	37
3.3.1) <i>Transkription</i>	37
3.3.2) <i>Codierung</i>	38
4.) ANALYSE DER FORSCHUNGSERGEBNISSE	40
4.1) FRAGEN (FORSCHUNGS-)ETHISCHER UND METHODISCHER NATUR	40
4.2) FRAGEN RECHTLICHER NATUR	45
4.2.1) <i>Urheberrecht</i>	46
4.2.2) <i>Daten-, Quellen- und Persönlichkeitsschutz</i>	48
4.2.3) <i>Anonymisierung und gestaffelte Zugänglichkeit durch die Einrichtungen</i>	50
4.2.4) <i>Informierte Einwilligung</i>	56
4.2.5) <i>Zwischenfazit ethische und rechtliche Fragen</i>	60
4.3) PRAKTISCH-TECHNISCHE PROBLEME IM UMGANG MIT QUALITATIVEN SOZIALWISSENSCHAFTLICHEN FORSCHUNGSDATEN	61
4.3.1) <i>Erschließung der Forschungsdaten</i>	64
4.3.2) <i>Auffindbarkeit der Forschungsdaten</i>	71
4.3.3) <i>Kooperationen mit den Fachcommunities</i>	73
4.3.4) <i>Datenauswahl und Arbeitsaufwand</i>	74
4.3.5) <i>Zwischenfazit praktisch-technische Probleme im Umgang mit qualitativen sozialwissenschaftlichen Forschungsdaten</i>	79
4.4) NACHNUTZUNG QUALITATIVER SOZIALWISSENSCHAFTLICHER FORSCHUNGSDATEN	81
4.4.1) <i>Konkreter Ablauf für die Nachnutzung</i>	83
4.4.2) <i>Zwischenfazit Nachnutzung qualitativer sozialwissenschaftlicher Forschungsdaten</i>	88
4.5) HERAUSFORDERUNGEN UND HINDERNISSE IN BEZUG AUF DAS MANAGEMENT QUALITATIVER SOZIALWISSENSCHAFTLICHER FORSCHUNGSDATEN	89
4.5.1) <i>Fachwissenschaftlicher Hintergrund und Ausbildung der Datenkurator_innen</i>	92
4.5.2) <i>Notwendigkeit der dauerhaften Finanzierung</i>	94
4.5.3) <i>Materialauswahl und -auffindbarkeit</i>	95
4.5.4) <i>Zwischenfazit hinsichtlich der Herausforderungen und Hindernisse in Bezug auf das Management qualitativer sozialwissenschaftlicher Forschungsdaten</i>	95
5.) FAZIT UND AUSBLICK	97
6.) ABKÜRZUNGSVERZEICHNIS	100
7.) LITERATUR- UND QUELLENVERZEICHNIS	101
7.1) LITERATURVERZEICHNIS	101
7.2) QUELLENVERZEICHNIS	118

8.) ANHANG	119
8.1) EXEMPLARISCHE LEITFADENÜBERSICHT.....	119
8.2) TRANSKRIPTE DER LEITFADENINTERVIEWS ZUR STUDIE	122

1.) Einleitung

1.1) Problematisierung und Fragestellung

Das Internet und bessere Computertechnik haben nicht nur neue Forschungsfragen ermöglicht, sondern auch für die Daten, die im Forschungsprozess gewonnen werden, neuartige Anwendungsfelder geschaffen (vgl. OECD 2007: 9). Diese technischen Innovationen haben einen veränderten Umgang mit Daten notwendig gemacht und zugleich deren Austausch in dem heute praktizierten Maße ermöglicht.¹ So schreiben Büttner et al.: „Die technischen Möglichkeiten sind vorhanden, nicht nur nie gekannte Mengen an Daten zu produzieren, sondern sie auch in neuer Weise (nach) zu nutzen, neu zu kombinieren und neuen Hypothesen zu unterziehen.“ (Büttner et al. 2011: 7) Wissenschaftliche Daten können nunmehr umfassend jenseits vom ursprünglichen Erhebungskontext verbreitet, zugänglich und für andere Projekte nutzbar gemacht werden.

Der Austausch und die (Nach-)Nutzung von Forschungsdaten sind seit Anfang der 2000er-Jahre verstärkt ein wissenschafts(-politisches) Thema. Internationale Gremien erwarten mehr Transparenz in der Forschung und eine stärkere Wiederverwendung von Daten, um den Mehrwert von Wissenschaftsinvestitionen zu maximieren. Forschungsförderinstitutionen wünschen einen offenen Zugang zu hochwertigen dokumentierten Forschungsdaten; akademische Verlage fordern Daten zwecks Überprüfung und Untermauerung von Ergebnissen an. Konzepte wie Open Science und Open Data, die den partizipativen freien Zugang zu Wissenschaft und Forschungsergebnissen propagieren, finden zunehmend Verbreitung.

Die Royal Society definiert Open Data exemplarisch folgendermaßen: „Open data requires effective communication through a more intelligent openness: data must be accessible and readily located; they must be intelligible to those who wish to scrutinise them; data must be assessable so that judgements can be made about their reliability and the

¹ Vgl. hierzu beispielsweise die „Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen“ aus dem Jahr 2003: „Mit dem Internet ist zum ersten Mal die Möglichkeit einer umfassenden und interaktiven Repräsentation des menschlichen Wissens, einschließlich des kulturellen Erbes, bei gleichzeitiger Gewährleistung eines weltweiten Zugangs gegeben.“ (ebd.: 1)

competence of those who created them; and they must be usable by others.“ (The Royal Society 2012: 9)²

Als paradigmatisch kann in diesem Zusammenhang die Aussage der ehemaligen Vizepräsidentin der Europäischen Kommission Neelie Kroes gelten, die 2012 für die Digitale Agenda zuständig war: „Die Steuerzahler sollten nicht zweimal für Forschungsergebnisse zahlen müssen, und sie müssen problemlos auf Rohdaten zugreifen können. Wir wollen bei der Verbreitung und Nutzung von Forschungsergebnissen die nächste Stufe erreichen. Daten sind das neue Erdöl.“ (EK 2012: 2) Fecher und Friesike (2013: 1-6) unterscheiden in diesem Kontext fünf verbreitete Ansätze und Motivationen: ein Ansatz, der Wissensproduktion verbessern will, einer der Wissenschaft öffentlich zugänglich machen will, dann derjenige, der sich Demokratisierung durch einen breiteren Zugang zu Wissen verspricht, ein infrastruktureller Ansatz und zuletzt eine Perspektive, die versucht andere Verfahren zur wissenschaftlichen Qualitätsmessung zu entwickeln.

Beschäftigten sich zunächst eher die Fächer, die viele (quantitative) Messdaten produzieren, damit, wie mit den von ihnen geschaffenen und verarbeiteten Daten verfahren werden soll, sind diese Fragen mittlerweile flächendeckend in fast allen Wissenschaftsdisziplinen von Relevanz (vgl. Harbeck/Kaun 2019: 2). Das Management quantitativer Forschungsdaten ist etabliert und kann sich auf eine mitunter langjährige Praxis beziehen, wohingegen „die Etablierung geeigneter Infrastrukturen für qualitativ forschende Disziplinen noch in den Kinderschuhen“ (Sterzer/Kretzer 2019: 110) steckt. In den Empfehlungen des Wissenschaftsrats bezüglich der Infrastrukturentwicklung der deutschen Wissenschaftslandschaft wird die Einrichtung von Forschungsdateninfrastruktureinrichtungen mit der Spezialisierung auf qualitative Daten als „besonders erforderlich“ (WR 2012: 11) beschrieben.³

Es besteht eine Diskrepanz zwischen Implementierungswünschen und mangelnder Infrastruktur, insbesondere weil Universallösungen für den Umgang mit qualitativen sozialwissenschaftlichen Forschungsdaten selten ausreichen (vgl. Harbeck/Kaun 2019: 4). Eine Übertragung von Datenmanagementmethoden und Infrastrukturlösungen von

² Vgl. auch: „Research data should be part of the scientific commonwealth, in the same way ideas are shared. For the sake of open scientific discourse and scholarship, data should be treated like fresh air freely available to all. Treating research data as property will ultimately restrict the interpretation and development of ideas. Science that relies on the ownership of facts and data seeks only to generate commercial value -- and is a very narrow view of science.“ (Jacobs/Humphrey 2004: 28)

³ Corti und Fielding haben im Dezember 2015 eine Zählung beim Datenrepositorien-Register re3data durchgeführt und kamen dabei auf ca. 1400 Datenrepositorien, wobei sich 378 von diesen auf Daten aus den Geistes- und Sozialwissenschaften spezialisiert hatten (vgl. Corti/Fielding 2016: 3). Auf persönliche Nachfrage via E-Mail teilte mir Gabriele Kloska für das re3data Editorial-Team mit, dass am 27.1.2020 2455 Repositorien auf re3data gelistet sind, wobei 819 davon aus dem Bereich 'Humanities and Social Sciences' und 393 aus den 'Social and Behavioural Sciences' stammen.

quantitativen Daten auf qualitative Daten ist nicht unmittelbar möglich. Das „Datenmanagement als vorzugsweise technische Disziplin zu verstehen greift viel zu kurz“ (Büttner et al. 2011: 21), die Charakteristika und (theoretisch-methodischen) Fragestellungen erfordern vielfach spezifische Lösungen.

Diese Studie befasst sich in der Hauptsache mit qualitativen sozialwissenschaftlichen Forschungsdaten, die in den ethnologischen Fächern, der Soziologie sowie Teilen der Medien-, Geschichts-, Geschichts- und Politikwissenschaften und der Psychologie erzeugt werden. Innerhalb dieser Disziplinen gibt es Unterschiede bezüglich deren Methoden und Umgang mit Daten, wobei diese Studie darauf abzielt deren Gemeinsamkeiten im Forschungsdatenmanagementprozess herauszuarbeiten.⁴ Fachliche Spezifika werden ebenfalls dargestellt, selbst wenn diese nicht immer für jedes Feld der qualitativen sozialwissenschaftlichen Forschungsdaten relevant sind. Hierbei ist jedoch nicht der gesamte Prozess des Forschungsdatenmanagements gleichermaßen von Bedeutung, vielmehr wird ein besonderes Augenmerk auf die Archivierung der Daten sowie deren Nachnutzungspotenzial gelegt.

Neben schriftlichen Quellen basiert diese Untersuchung auf leitfadengestützten Interviews. Zu diesem Zweck wurden sieben Gespräche mit Expert_innen aus Datenarchiven geführt, die anonymisiert aufgezeichnet wurden. Auf eine explizite Nennung der Einrichtung als auch auf die Namen der Interviewteilnehmer_innen ist, auf Wunsch der Teilnehmenden, verzichtet worden.

Das übergreifende Erkenntnisinteresse dieser Studie zielt darauf den gegenwärtigen Zustand der Archivierung und Nachnutzung von qualitativen Forschungsdaten, samt seiner Probleme und Potentiale, zu beleuchten. Sie zeigt auf welche Rolle Infrastruktureinrichtungen, vor dem Hintergrund von (forschungs-)ethischer und methodischer, rechtlicher und praktisch-technischer Probleme, dabei einnehmen können. Gleichzeitig stellt sie dar welche Aufgaben bezüglich der Auswahl, der Erschließung und der Nachnutzung übernommen werden müssen. Daraus ergeben sich folgende Themenkomplexe und Fragen, die gleichzeitig zur Strukturierung der Studie genutzt werden:

⁴ An manchen Stellen liegt der Schwerpunkt, begründet durch die fachliche Aus- und Vorbildung des Autors, insbesondere auf den Ethnologien.

1.) Zunächst wird auf allgemeiner Ebene geklärt, was Forschungsdaten sind und welche Positionen die verschiedenen Akteursgruppen, die in diesem Feld relevant sind, dahingehend einnehmen. Hierbei werden wissenschaftspolitische Akteur_innen, wie Förderorganisationen sowie wissenschaftliche Communities in Form von Fachgesellschaften und Forschungscommunities betrachtet. Die Standpunkte dieser Gruppen sind von zentraler Bedeutung, da diese jeweils eigene, mitunter konflikthafte Interessen und Bedürfnisse besitzen, die andauernde Aushandlungen bedürfen. Die Datenarchive sind auf diese Wissenschaftsnetzwerke und deren fachliche Impulse angewiesen, da sie ihre Angebote nicht losgelöst von der Wissenschaftslandschaft entwickeln können. Sie müssen vielmehr mit ihren Angeboten auf deren Anforderungen reagieren.

2.) Im Anschluss wird die Spezifik qualitativer sozialwissenschaftlicher Forschungsdaten beschrieben und aufgezeigt, was sie für das Forschungsdatenmanagement bedeutet. Schlussendlich soll auf Grundlage von Publikationen (u. a. Statements von Wissenschaftler_innen und Fachgesellschaften) herausgearbeitet werden, welche Hemmnisse bei der Etablierung von (qualitativen) Forschungsdatenmanagementprozessen bestehen und welche Kritik an dieser Konzeption auf allgemeiner Ebene existiert.

3.) Erörtert wird, welche Arbeitsprozesse die Sammlung und Archivierung von qualitativen sozialwissenschaftlichen Forschungsdaten umfassen und welche Rollen Forscher_innen und die Infrastruktureinrichtungen dabei übernehmen. Welche zeitlichen Ressourcen für die Aufbereitung der Daten aufgewandt werden müssen und inwiefern sich der Umgang mit den Forschungsdaten in den eigentlichen Forschungsprozess integrieren lässt? Konkrete Entwicklungen und insbesondere Probleme und Erfolge werden im Verlauf nachgezeichnet. Wie erschließen und kontextualisieren die Befragten die Forschungsdaten und stellen deren Auffindbarkeit sicher? Wie findet die Kenntnisweitergabe statt und welche Kooperationen zwischen den verschiedenen Beteiligten am Forschungsdatenmanagement existieren?

4.) Die Studie arbeitet heraus, welche Probleme sich auf (forschungs-)ethischer, praktischer und rechtlicher Ebene im Zuge der Archivierung und Nachnutzung von qualitativen sozialwissenschaftlichen Forschungsdaten ergeben und inwiefern die befragten Einrichtungen mit diesen Problemen umgehen.

5.) Außerdem wird der Frage nachgegangen werden, wie viel Arbeit in die Aufbereitung der Daten investiert wird und auf welcher Grundlage diese als archivierungswürdig ausgewählt werden?

6.) Die Nachnutzung der archivierten Daten ist ein zentrales Anliegen innerhalb des Forschungsdatenmanagementprozesses, deshalb wird dargestellt wie Nachnutzungsprozesse konkret ablaufen und welche Formen Nachnutzungen (potenziell) annehmen können.

7.) Schlussendlich arbeitet diese Studie, basierend auf den mit Infrastruktureinrichtungen geführten Interviews, heraus, welche spezifischen Herausforderungen und Hindernisse in Bezug auf das Management qualitativer sozialwissenschaftlicher Forschungsdaten existieren und was dies für die Entwicklung des Datenmanagements und dessen Praxis in den kommenden Jahren bedeutet.

1.2) Aufbau der Arbeit

Die vorliegende Studie gliedert sich in vier Abschnitte: Ein Erster definiert den Begriff der Forschungsdaten auf einer allgemeinen Ebene und stellt wissenschaftspolitische Positionen, Forderungen und Ziele, die mit dem Forschungsdatenmanagement verbunden sind, dar.

Ein weiterer Abschnitt erläutert das methodische Vorgehen und die, der Erhebung zugrundeliegende, Methode des Expert_inneninterviews samt Auswertung.

Im nächsten Teil folgt die Darstellung der Forschungsergebnisse, wobei auf die Themenbereiche Sammeln, Archivieren, Finden, Bereitstellen sowie die Nachnutzung unter (forschungs-)ethischen, rechtlichen, ökonomischen Gesichtspunkten eingegangen wird. Dies umfasst eine Betrachtung der Spezifik qualitativer Methoden und Forschungsdaten entlang der Achsen praktisch-technischer Probleme sowie Fragen hinsichtlich rechtlicher und (forschungs-)ethischer Aspekte. Unterschiede und Konflikte erschweren dabei eine passgenaue Übertragung der Forschungsdatenkonzeption aus dem Bereich der quantitativen Daten. Außerdem werden Hindernisse und Bedarfe im Zuge des Forschungsdatenmanagements dargestellt.

Schließlich werden die gewonnenen Ergebnisse in einem gesonderten Fazit samt Ausblick zusammengefasst.

2.) Forschungsdaten(-management): Definition und wissenschaftspolitische Positionierungen

Der folgende Abschnitt definiert den Begriff Forschungsdaten, arbeitet die Gründe für die Forderung nach Forschungsdatenmanagement sowie die Implementierungsgeschichte dieses Wissenschaftsparadigmas heraus. Danach wird die Spezifik qualitativer sozialwissenschaftlicher Daten in diesem Kontext aufgezeigt. Im Zuge dessen werden wissenschaftspolitische Positionspapiere sowie eine Kritik am Forschungsdatenparadigma aus fachwissenschaftlicher Sicht dargestellt.

2.1) *Definition von Forschungsdaten*

Was unter Forschungsdaten verstanden wird, ist abhängig von der wissenschaftlichen Disziplin, genutzten Methoden wie Experimenten, Messungen, Befragungen, Feld- und Quellenforschung sowie der Forschungsfrage (vgl. Kindling et al. 2013: 2; Kindling/Schirnbacher 2013: 130). Ebenso relevant ist der Forschungsfokus als auch inwieweit die Daten verarbeitet und analysiert worden sind, was Einfluss darauf hat, unter welchen Bedingungen Forscher_innen bereit sind, diese zu teilen.

Sie umfassen dabei ein Spektrum, das von den Daten zu Beginn eines Forschungsprozesses, über Arbeits- bzw. Intermediärdaten bis hin zu Ergebnisdaten bei Projektende reichen kann. „Die verschiedenen Wissenschaften sprechen von Messdaten, Rohdaten, empirischen Daten, Quelldaten, Forschungsrohdaten etc., verwenden viele unterschiedliche Begriffe und haben auch unterschiedliche Konzepte von dem, was für sie ein „Datum“ ist.“ (Büttner et al. 2011: 15) Forschungsdaten sind somit Quellenmaterialien, die für weitere Forschungsprozesse verwendet werden können, sowie Ergebnisdaten, aus denen Publikationen entstehen können (vgl. Degkwitz 2019: 7). Forschungsdaten erscheinen in diesem Fall als uneinheitlicher „Relationsbegriff“ (Queckbörner 2019: 10), wobei jeweils die Einbindung in eine Fachspezifik mit ihren distinkten Arbeitsabläufen sowie Methoden beachtet werden muss. Eine Anpassung an damit einhergehende Bedürfnisse ist erforderlich, „eine Übertragung auf andere Wissenschaftsdisziplinen nicht per se“ (Rümpel 2011: 32) möglich.

Der Umgang mit diesen Daten umfasst einen sogenannten Forschungsdatenmanagementprozess, der Aufgaben, wie das Verwalten, Organisieren und Verarbeiten der Daten beinhaltet. Diese Tätigkeiten werden oftmals in einem Datenmanagementplan festgehalten, der die systematische und dokumentierte Handhabung innerhalb des Projekts ermöglicht. Die notwendigen Schritte beginnend mit der Planung der Datenerhebung bis hin zur Nachnutzung werden oftmals mit Rückgriff auf das sogenannte DCC Curation Lifecycle Model idealtypisch entlang von acht Hauptschritten beschrieben: Konzeptionalisierung, Datenerstellung oder -übernahme, Bewertung und Auswahl, Aufnahme, Erhaltungsmaßnahmen, Speicherung, Zugriff sowie Nutzung und Wiederverwendung, Transformation (vgl. DDC: 2020). Hierbei sind die Anforderungen folgender Beteiligter in Einklang zu bringen: Primärforscher_innen, Beforschte, Forschungsförderungsorganisationen, Repositorien/Archive sowie Nachnutzende (vgl. DGfE 2017: 5 f.).⁵ Ziel ist, eine langfristige Nachnutzung und Bereitstellung von Daten durch zuverlässige Infrastrukturen (Forschungsdaten-Repositorien) sicherzustellen (vgl. Scholze et al. 2016: 157). Schönbrodt et al. nennen folgende Qualitätsmerkmale für ein verlässliches Repository:

„wirtschaftliche bzw. ideologische Unabhängigkeit und wissenschaftliche Professionalität [...] Persistenz der Daten [...] muss gewährleistet sein. Die Zugänglichkeit der Daten: Es muss möglich sein, die Daten öffentlich und kostenlos abzurufen; es muss aber auch möglich sein, Zugangsbeschränkungen [...] zu definieren. [...] Die Identifizierbarkeit der Daten: Es muss ein persistenter Identifikator [...] vergeben werden [...] Die Klärung der Rechte an den Daten [...] Die Möglichkeit, Dateien bzw. Datensätze sowohl öffentlich als auch nicht-öffentlich abzuspeichern.“ (2017: 22)

Diese Datenzentren und -archive sind entweder fachspezifisch oder fächerübergreifend (z. B. Zenodo) und bieten die technische Infrastruktur zur längerfristigen Datensicherung und Verfügbarmachung (vgl. Kindling et al. 2013: 20). Die Bereitstellung umfasst eine Datenaufbereitung, in Form von fachwissenschaftlicher Erschließung und Verbreitung, samt Metadatenvergabe sowie eine Datendokumentation, die Kontextinformationen beinhaltet, um so sicherzustellen, „dass Datenzugriff und -auswertung unabhängig vom Datenerzeuger

⁵ Zur Prozesshaftigkeit des Forschungsdatenmanagements vgl.: „Data archiving is a process, not an end state where data is simply turned over to a repository at the conclusion of a study. Rather, data archiving should begin early in a project and incorporate a schedule for depositing products over the course of a project's life cycle and the creation and preservation of accurate metadata, ensuring the usability of the research data itself. Such practices would incorporate archiving as part of the research method.“ (Jacobs/Humphrey 2004: 28)

möglich wird und bleibt.” (Büttner/Hobohm 2011: 14) Datenarchive können verschiedene Kurationsstätigkeiten durchführen, zu denen die Überprüfung und Validierung, die Gewährleistung der Einhaltung von Zustimmungs- und Vertraulichkeitsvereinbarungen, die Anonymisierung, wo erforderlich, und die längerfristige Speicherung von Daten gehören (vgl. Corti 2018: 21). Dauerhafte Nachnutzung bedarf neben der Archivierung auch Schnittstellen zum Abrufen der Daten und Recherchemöglichkeiten, beispielsweise via zentralen Katalogen, die fachspezifische Funktionalitäten aufweisen (vgl. Sahle/Kronenwett 2013: 93; Scholze et al. 2016: 158). Das Potenzial für Sekundärnutzungen unterscheidet sich ebenfalls zwischen den verschiedenen Datensätzen. Außerdem können, je nach Kontext der Rezeption, dieselben Daten andere Erkenntnisse ermöglichen (vgl. Voß 2013: 8). Borgman beschreibt das sich verändernde Nutzungspotential folgendermaßen: „Some types of data have both immediate and enduring value, some gain value over time, some have transient value, and yet others are easier to recreate than to curate.“ (Borgman 2010: 3)

2.2) Allgemeine wissenschaftspolitische Positionierungen

Vonseiten der Wissenschaftsförderung sind Forschungsdaten in Deutschland seit etwas mehr als zwanzig Jahren eine bedeutsame Fragestellung. Zu Beginn lag der Fokus auf der Nachprüfbarkeit und erst später hat das Thema der Nachnutzbarkeit der Daten als wissenschaftliche Ressource an Relevanz gewonnen. Disziplinübergreifend sollen diese Daten nicht mehr nur gespeichert werden, sondern möglichst auch frei austauschbar und verfügbar zur Nachnutzung sein. Argumente dafür können entlang zweier Achsen betrachtet werden: einerseits der Anspruch der Öffentlichkeit, Zugang zu Daten zu bekommen, die von ihr finanziert wurden, andererseits der Wunsch der Wissenschaft bereits erhobene Daten nutzen zu können und getrieben vom Interesse von Datenproduzent_innen hin zu dem von potenziellen Datennutzer_innen. Vier zentrale Gründe werden von Autor_innenseite genannt (vgl. Borgman 2010: 14; Harbeck/Kaun 2019: 2):

- 1.) Die Ergebnisse öffentlich finanzierter Forschung sollen der Öffentlichkeit zugänglich gemacht werden, um ein breiteres Maß an interdisziplinärer Kollaboration sowie eine effektivere Bezugnahme auf Studien zu bewirken. Aufwendige Forschungswiederholungen und Überforschungen können vermieden werden und eine größere (Kosten-)Effizienz erreicht werden (vgl. Hollstein/Strübing 2018b: 2). Im Zuge dessen zeigt sich, dass „zugleich eine Neubestimmung der Rolle

von Wissenschaft in der Gesellschaft stattfindet“ (Eberhard/Kraus 2018: 42 f.), wobei die Wissenschaft, durch ihre Arbeit, einen Mehrwert für die sie finanzierende Öffentlichkeit hervorbringen soll. Einerseits lässt sich eine Ökonomisierung der Wissenschaftslandschaft und Kommodifizierung von Forschungsergebnissen konstatieren (vgl. Pels 2018a: 391 ff.), zugleich existieren demokratisierende Elemente im Hinblick auf freie Zugänglichkeit und Verfügbarkeit: „Im Sinne von Open Science soll auch das Versprechen von mehr Teilhabe, Gerechtigkeit und Innovation eingelöst werden.“ (Imeri 2018b: 217)

- 2.) Für Nutzer_innen birgt es das Potenzial, andere interessierte Forscher_innen in die Lage zu versetzen, Daten wiederzuverwenden und neue Fragen bezüglich vorhandener Daten in Form von Sekundäranalysen zu stellen. So sind umfassende Vergleiche, historische Erhebungen, Langzeitstudien sowie Metaanalysen möglich (vgl. Hollstein/Strübing 2018b: 2). Besonders wertvoll sind diese Daten, wenn sie sich nicht unmittelbar replizieren lassen z. B. bei singulären Ereignissen. Der Gefahr des Datenverlustes durch technische Probleme, den Wechsel von Mitarbeiter_innen in Forschungsprojekten, kann so begegnet und oft nur einmalig und selten erschöpfend analysierte Datensammlungen genutzt werden (vgl. Smioski 2011: 220).
- 3.) Forschungsdatenreferenzierung kann helfen, die wissenschaftliche Praxis und die Forschungsqualität zu verbessern. Fremde Forschungsdaten können genutzt werden, um methodische Zugänge zu reflektieren, beispielsweise durch die Verwendung der Daten zu Schulungszwecken (vgl. Smioski 2011: 222).
- 4.) Forschungsergebnisse können reproduziert sowie reanalysiert werden und der Forschungsprozess transparenter werden, was eine Qualitätssicherung bzw. eine Überprüfbarkeit von Ergebnissen ermöglicht (vgl. Hollstein/Strübing 2018b: 2).

Während in Großbritannien schon seit den 1970ern von Seiten des United Kingdom's Economic and Social Research Council und dessen Vorgängerorganisation Forschungsdatenregularien existierten, die vorsahen, dass die gewonnenen quantitativen Daten (ab 1995 auch qualitativen Daten) formal archiviert werden (vgl. Bishop/Kuula-Luumi 2017: 2) veröffentlichte die DFG erstmals im Jahr 1998 Äußerungen zu diesem Thema. Die Denkschrift mit dem Titel „Sicherung guter wissenschaftlicher Praxis“, wurde

in Reaktion auf einige skandalträchtige Fälle von wissenschaftlichem Fehlverhalten publiziert. So war aufgedeckt worden, dass die renommierten biomedizinischen Forscher_innen Friedhelm Hermann und Marion Brach im Zeitraum von 1988 bis 1996 während ihrer Forschungstätigkeit an deutschen Universitäten ihre Forschungsergebnisse stark verfälscht und manipuliert hatten (vgl. Koenig 1997; Finetti/Himmelrath 1999: 33-61).⁶ Neben der Fälschung von Experimenten und Plagiaten nennen Corti und Fielding das sogenannte „Schreibtischschubladen“-Problem, wo negative bzw. zweideutige Ergebnisse weniger wahrscheinlich veröffentlicht werden. War das Aufspüren solcher Probleme im analogen Wissenschaftsbetrieb noch mühselig, hat die Digitalisierung von Zeitschrifteninhalten mittlerweile eine Massenvvalidierungsanalyse von Artikeln ermöglicht, was zu stark angestiegenen Rückzugsraten geführt hat (vgl. Corti/Fielding 2016: 8). Die Empfehlungen der DFG forderten die sichere Aufbewahrung von Primärdaten, die als Basis für Veröffentlichungen gedient haben, in der datenerzeugenden geförderten Einrichtung für zehn Jahre (vgl. DFG 2013 [1998]: 21). Kontrolle und Nachprüfbarkeit von Forschungsergebnissen waren die zentralen Beweggründe, das Thema Nachnutzung von Daten spielte noch keine explizite Rolle (vgl. DFG 2013 [1998]: 13, 22). Die ausgesprochenen Empfehlungen wurden später in ähnlicher Form auch von weiteren Wissenschaftsorganisationen adaptiert (vgl. Klump 2010: 104).

In den folgenden Jahren publizierten die DFG, aber auch andere Forschungsförderungseinrichtungen und Wissenschaftszusammenschlüsse weitere Verlautbarungen dazu, wobei das Thema Datennachnutzung zunehmend wichtiger wurde.⁷ Diese nichtfachspezifischen Statements stellten Ansprüche an die Wissenschaftscommunity, die auf diese Anforderungen nunmehr reagieren sollte.

⁶ Für weitere Fälle siehe das Kapitel „Nach dem Paukenschlag - Jüngste Fälschungsfälle und -vorwürfe“ in Finetti/Himmelrath 1999: 124-144.

⁷ Auf die Vielzahl der Texte kann hier aus Platzgründen nur exemplarisch eingegangen werden: Neben den angeführten Stellungnahmen sei noch auf folgende Weitere hingewiesen: 2010 veröffentlichte die Allianz der deutschen Wissenschaftsorganisationen ihre „Grundsätze zum Umgang mit Forschungsdaten“, wonach Forschungsdaten nicht mehr nur zur Überprüfung dienen sollten, sondern „in hohem Maße auch der Erzielung künftiger Ergebnisse [...] Mit dem Ziel, die Qualität, Produktivität und Wettbewerbsfähigkeit der Wissenschaft zu fördern“ (AddW 2010: 5). Der Wissenschaftsrat (WR) veröffentlichte 2012 seine „Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020“ in denen er unter anderem forderte, dass die „Erhebung und Publikation von Forschungsdaten [...] als eigenständige Forschungsleistung anerkannt“ werden sollte (WR 2012: 14). Zudem kann auf Empfehlungen von Hochschuleseite verwiesen werden, siehe beispielhaft hierzu die Publikation der Hochschulrektorenkonferenz (HRK) „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können. Orientierungspfade, Handlungsoptionen, Szenarien“ (HRK 2015). Neben diesen bundesweiten Bestrebungen gibt es noch zahlreiche Stellungnahmen und Projekte auf Länderebene, exemplarisch soll hier auf die „Open-Access-Strategie für Berlin“ hingewiesen werden (vgl. Arbeitsgruppe Open-Access-Strategie Berlin 2015: 17-19).

In den „Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten“ betont die DFG den finanziellen Aufwand der Datenerhebung und fordert den freien, öffentlichen und nachhaltigen Zugang zu diesen Daten, um diese für neue Forschungsfragen aber auch zu Prüfzwecken nutzen zu können (vgl. DFG 2009: 1).

So schreibt sie:

„In jedem Fall sollten die erhobenen Daten nach Abschluss der Forschungen öffentlich zugänglich und frei verfügbar sein. Dieses ist die wesentliche Voraussetzung dafür, dass Daten im Rahmen neuer Fragestellungen wieder genutzt werden können sowie dafür, dass im Falle von Zweifeln an der Publikation die Daten für die Überprüfung der publizierten Ergebnisse herangezogen werden können.“ (DFG 2009: 1)

In ihrem „Leitfaden für die Antragsstellung“ fordert die DFG die von ihr geförderten Wissenschaftler_innen zudem auf, darzustellen, inwiefern für die von ihnen gewonnen Forschungsdaten eine Nachnutzung möglich ist und wie diese anderen Wissenschaftler_innen, unter Rückgriff auf fachspezifische Standards, zur Verfügung gestellt werden können (vgl. DFG 2018: 5). Die „Leitlinien zum Umgang mit Forschungsdaten“ enthalten schließlich die Forderung, dass die Forschungsdaten schnellstmöglich zur Verfügung stehen sollen, falls Rechte Dritter die Zugänglichmachung nicht einschränken (vgl. DFG 2015: 1). Die Bereitstellung von Forschungsdaten sollte zudem bei wissenschaftlicher Leistungsbeurteilung in Zukunft mehr anerkannt werden (vgl. DFG 2015: 2).

Auch die „Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen“, die Forderungen nach einer freier zugänglichen Wissenschafts- und Informationspraxis in Form von Open-Access-Publikationen umfasst, behandelte das Thema öffentlich finanzierter Forschungsdaten. Open-Access-Veröffentlichungen sollten „originäre wissenschaftliche Forschungsergebnisse ebenso wie Ursprungsdaten, Metadaten, Quellenmaterial, digitale Darstellungen von Bild- und Grafik-Material und wissenschaftliches Material in multimedialer Form“ umfassen (Berliner Erklärung 2003: 1).

Auf europäischer Ebene veröffentlichte die OECD Empfehlungen, worin sie koordinierte Maßnahmen auf nationaler sowie internationaler Ebene verlangte, um den Zugang zu Daten aus öffentlicher Förderung zu erweitern. Dies habe erhebliche Vorteile für Wissenschaft, Wirtschaft und Gesellschaft, würde innovationsfördernd wirken und damit die Qualität und

Produktivität des Wissenschaftssystems weltweit steigern. Unangemessene Zugangsbeschränkungen bergen hingegen das Risiko der Beeinträchtigung der Qualität und Effizienz der wissenschaftlichen Forschung (vgl. OECD 2004).⁸

Die Europäische Kommission hatte ebenfalls im Zuge von „Horizon 2020“ die öffentliche Bereitstellung und Zugänglichmachung von Forschungsdaten als Pilotprojekt zum Ziel. Die zwei Schritte des sogenannten „Open Research Data Pilot“ umfassen dabei: „Step 1 - they must deposit the research data described above, preferably in a research data repository“ und „Step 2 – as far as possible, projects must then take measures to enable third parties to access, mine, exploit, reproduce and disseminate (free of charge for any user) this research data.“ (EC 2017: 10) Begründet wurde dies entlang von vier Argumenten: die Möglichkeit auf früheren Forschungsergebnissen aufzubauen (verbesserte Qualität der Ergebnisse), die Zusammenarbeit zu fördern und Doppelarbeit zu vermeiden (größere Effizienz), die Innovation zu beschleunigen und Bürger_innen und Gesellschaft einzubeziehen (erhöhte Transparenz des wissenschaftlichen Prozesses) (vgl. EC 2017: 5).

Die „FAIR Guiding Principles“ sollen als Modell dienen, das darauf abzielt, bei Forschungsdaten, die Auffindbarkeit (Findability), Zugänglichkeit (Accessibility), Interoperabilität (Interoperability) und Nachnutzbarkeit (Reusability) für Menschen und Maschinen sicherzustellen (vgl. Wilkinson et al. 2016: 3). Diese Prinzipien sollen Datenproduzent_innen in ihrer Arbeit unterstützen, Nachnutzung vereinfachen und so maximalen Nutzen für eine möglichst zugängliche Wissenschaft generieren (ebd.: 1). Die FAIR-Prinzipien stellen in gewisser Weise eine Abkehr vom Verständnis vollkommen frei zugänglicher Daten dar. Es gilt nach wie vor, die Forschungsdaten nachnutzbar zu machen. Das „Konzept trägt aber dem Umstand Rechnung, dass Offenheit eben Unterschiedliches bedeuten kann oder muss. Transparent gemacht werden damit vor allem die Bedingungen, unter denen Daten nutzbar sind.“ (Imeri 2019: 49)⁹ Bisher gibt es jedoch keine einheitliche nationale Politik in den OECD-Ländern, die den Datenaustausch auf diese Weise vorschreibt, obwohl die Anforderungen in Bezug auf Datenmanagementpläne steigen, in

⁸ Vgl. auch: „Besides, access to research data increases the returns from public investment in this area; reinforces open scientific inquiry; encourages diversity of studies and opinion; promotes new areas of work and enables the exploration of topics not envisioned by the initial investigators.“ (OECD 2007: 3)

⁹ Vgl. hier die Aussagen der Europäischen Kommission im Zuge von „Horizon 2020“: „The Commission's approach can be described as ‘as open as possible, as closed as necessary.’“ (EC 2017: 8)

denen die Forschenden aufgefordert werden, den Datenaustausch und die Kuratierung der Daten zu berücksichtigen (vgl. Bishop/Kuula-Luumi 2017: 2).

Eine Ergänzung zu diesen Prinzipien stellen die sogenannten CARE-Grundsätze für die Verwaltung indigener Daten dar. Sie wurden erstellt, da vorhandene Prinzipien Machtunterschiede und historische Zusammenhänge im Datenmanagement nicht thematisieren (vgl. RDAIIDSIG 2019: 1).

Sie umfassen vier Bereiche:

1.) Collective Benefit: Die Datenökosysteme sollen so gestaltet sein und funktionieren, dass es indigenen Gruppen ermöglicht wird aus den Daten Nutzen ziehen zu können (ebd.: 2).

2.) Authority to Control: Die Rechte und Interessen indigener Gruppen an indigenen Daten müssen anerkannt und ihre Autorität zur Kontrolle dieser Daten gestärkt werden (ebd.: 3).

3.) Responsibility: Diejenigen, die mit indigenen Daten arbeiten, haben die Verantwortung, die Art und Weise zu erläutern, wie diese Daten zur Unterstützung der Selbstbestimmung und des kollektiven Nutzens der indigenen Gruppen verwendet werden können (ebd.: 4).

4.) Ethics: Die Rechte und das Wohlergehen der indigenen Gruppen sollten in allen Phasen des Datenlebenszyklus und im gesamten Datenökosystem das vorrangige Anliegen sein (ebd.: 5).

2.3) Spezifik qualitativer sozialwissenschaftlicher Daten und deren Nachnutzung

Der im Zuge des Forschungsdatenmanagements verwendete Datenbegriff ist mitunter zu generisch, deshalb benötigt er bezüglich qualitativer sozialwissenschaftlicher Daten eine Spezifizierung. Qualitative Forschung unterscheidet sich sowohl durch Forschungsintention als auch durch die Forschungsmethoden von quantitativen Verfahren. Die Diskussion um qualitative sozialwissenschaftliche Forschungsdaten war lange durch den Fokus auf qualitative Interviews geprägt soll aber nun ebenfalls verstärkt anderes qualitatives

Datenmaterial miteinbeziehen.¹⁰ Das Bandbreite qualitativer sozialwissenschaftlicher Daten und deren Erhebungsmethoden werden im Folgenden dargestellt.

Ein quantitativer Zugang möchte Sachverhalte durch Zählung und Messung via mathematisch-statistischer Methoden erforschen. Das zentrale Forschungsinstrument stellen hierbei schriftliche oder mündliche Befragungen nach einem standardisierten Raster dar, die stark strukturierte Daten erzeugen. Daten werden hier vielfach noch als von Umwelteinflüssen bereinigte Fakten verstanden, die über beobachterunabhängige Objektivität verfügen. Sie lassen sich reproduzieren und als Datenobjekte oft losgelöst vom Kontext der Generierung austauschen (vgl. Ballsun-Stanton 2012: 291; Reichert 2018: 18 ff.).

Qualitative Forschung zielt darauf ab, durch alltagsnahe Beobachtung Phänomene einer Beschreibung und Analyse zu unterziehen, um so Denkweisen und Handlungen von Individuen und Gruppen zu verstehen. Hier gelten Daten eher als Aufzeichnungen von subjektiven Beobachtungen, die sinngebende Interpretation, Filterung und Kontextualisierung benötigen, um zu bestimmen, was für den Forschungsprozess von Relevanz ist (vgl. Ballsun-Stanton 2012: 18). Die Daten lassen sich dabei in einem Verarbeitungskontinuum von Rohdaten bzw. Ausgangsdaten bis hin zu prozessierten und eingearbeiteten (Ergebnis-)Daten in Publikationen verorten (vgl. Huber 2019: 4).

Wobei Imeri mit Bezug auf die Ethnologien darauf hinweist, „dass die in der Debatte um Forschungsdaten regelmäßig anzutreffenden Unterscheidungen von (objektiven) Rohdaten, Forschungsprimärdaten, prozessierten Daten und Interpretationen, die am Ende der Datenanalyse stehen, nicht ohne Weiteres auf die gesamte Forschung der ethnologischen Fächer übertragen werden können.“ (Imeri 2018b: 228) Teilweise wird der Datenbegriff vermieden und eher von Unterlagen, Dokumenten und Materialien gesprochen (vgl. Imeri 2018a: 72). Ergänzend kommentieren Pels et al., dass die Grenzziehung auch immer von den Bedürfnissen potenzieller Nutzer_innen abhängig ist: „Moreover, the boundaries between raw and processed data are drawn differently, depending on whether the audience in question is interested in mere verification, publication or reuse, and the drawing of that boundary, and

¹⁰ Vgl. die zentralen Ergebnisse des interdisziplinären Workshops zu „Archivierung und Zugang zu qualitativen Daten“, des RatSWD: „Die Diskussion in den Sozialwissenschaften hat sich bislang besonders auf qualitative Interviews bezogen, sollte jedoch zukünftig stärker auch andere qualitativ-empirische Materialien (ethnografische Daten, Bild- und Videomaterial) einbeziehen und die datenschutzrechtlichen und forschungsethischen Probleme sowie mögliche Nachnutzungspotenziale adressieren.“ (Hollstein/Strübing 2018c: 101)

therefore the provision of access to the data concerned, is the ethical responsibility of the researcher.“ (2018b: 399)

Qualitative sozialwissenschaftliche Daten unterscheiden sich von quantitativen „durch eine höhere Vielfalt hinsichtlich ihrer Medialität, Materialität und Zeichenform, ihrer kontextuellen Spezifität und ihres höheren Interpretationsbedarfs deutlich.“ (Knoblauch/Wilke 2018: 52) Qualitative Methoden umfassen einen breit gefächerten und nicht standardisierten Methodenmix, der unterschiedliche Beobachtungsarten, Einzelinterviews sowie Gruppendiskussionen umfasst (vgl. Harbeck et al. 2018: 127; Imeri et al. 2019: 10). Corti beschreibt für die verschiedenen Forschungsdisziplinen unterschiedliche Vorgehensweisen:

„Disciplines like sociology, anthropology and education tend to prefer interviewing and ethnographic methods, generating data such as interview or focus group recordings and transcripts, diaries, other kind of observations, fieldnotes and kinship diagrams. Political science qualitative methods prefer to focus on text analysis using historical and legal documents, newspaper sources, political manifestoes and records of meetings, for example, in the policy environment or in human rights organisations. Humanities scholars such as historians or linguists often focus on a close examination of the source, using methods of text or audio-visual analysis to discover linguistic features of speech and on verbal communication and create transcribed and annotated text and audiovisual data.“ (2018: 15)

Als analoge und digitale Forschungsdaten beinhalten sie beispielsweise Notizen, Protokolle, Transkripte, Feldtagebücher und andere Dokumente, sowie audio-visuelle Daten wie Audioaufnahmen, Fotos und Videos. Manche dieser Daten werden nicht im Feld erzeugt, sondern sind in diesem in Form von Gegenständen oder Dokumenten bereits vorhanden (vgl. Meier zu Verl/Meyer 2018: 83; Smioski 2011: 226).

Eine besonders in den Ethnologien weitverbreitetes Forschungsszenario stellt die Feldforschung durch teilnehmende Beobachtung dar, die auf der Idee beruht, „dass die Beobachtung nicht vom Miterleben, von der Teilnahme am Alltag der Gesellschaft abgetrennt werden darf.“ (Hahn 2014: 73) In der Feldforschung wird jene Methode noch durch Interviews, Archivrecherchen und Kartierungsverfahren ergänzt. Dieser methodische Pluralismus erzeugt einen zumeist qualitativen Forschungsdatenmix, in dem die unterschiedlichen Daten wechselseitig aufeinander Bezug nehmen (vgl. Imeri 2017: 168). Die Forschungen sind in der Regel durch die Situations- und Beobachter_innenabhängigkeit

nicht wiederholbar, standardisierbar und wenig planbar, was eine ressourcenintensive Flexibilität zum Beispiel in Hinblick auf die Methoden erfordert (vgl. Harbeck et al. 2018: 128).

Um qualitative Forschung zu realisieren, ist vielfach ein umfangreicher Beziehungsaufbau zwischen Forschenden und Beforschten notwendig. „In anthropology and ethnography (but this may be valid for social science in general), methods and ethics are mutually supportive, congruent and sometimes even identical. They rest on the same epistemological foundation of a process of mutual learning that builds social relationships on varying degrees of trust.“ (Pels 2018a: 394) Da während der Forschung oftmals Einblicke in sehr persönliche Lebensbereiche, die mitunter für Akteur_innen gefährlich sein können, gewährt werden, muss dieser Feldzugang oft erst durch vertrauensbildende Maßnahmen erarbeitet werden (vgl. Breidenstein et al. 2013: 60 ff.). Anders als bei großen quantitativen Studien wie beispielsweise dem Zensus ist die dauerhafte Zusammenarbeit und Billigung der Akteur_innen essenziell, weil die Forscher_innen von den Forschungssubjekten abhängig sind (vgl. Imeri 2017: 169). In den Ethnologien wird zudem von Forschendenseite diskutiert, inwiefern die Forschungssubjekte durch die notwendige Kooperationsbeziehung als Ko-Produzent_innen ebenfalls zur Datengenerierung beitragen.

Für das Forschungsdatenmanagement ergeben sich relevante Fragen hinsichtlich des Forschungsbeginns und des -abschlusses. Insbesondere bei Feldforschungen, wo Zugang zum und Aufbau von Beziehungen im Feld mitunter sehr arbeitsintensiv sind und die Orte oftmals über einen längeren Zeitraum mehrfach besucht werden, wird das gewonnene Material von Forschenden über eine sehr lange Zeitspanne und nicht nur über einen begrenzten Projektzeitraum genutzt. Der Zeitpunkt des tatsächlichen Forschungsabschlusses an dem die Daten in ein Repository weitergegeben werden können ist deshalb unklarer zu bestimmen als bei anderen Forschungsvorhaben (vgl. Imeri et al. 2019: 12).

Eine umfassende Objektivierung der Forschungserfahrungen ist ebenfalls schwierig, da sich diese vielfach „auf der Ebene präreflexiver Vertrautheit im Gedächtnis und Körper“ der Forschenden niederschlagen, was nur „situativ abrufbar ist.“ (Meier zu Verl/Meyer 2018: 83) Durch die spezifischen Erhebungsmethoden in Form der „sinnliche(n) Unmittelbarkeit der gesuchten Forschungserfahrung“ (Breidenstein et al. 2013: 33) gibt es bei qualitativen sozialwissenschaftlichen Daten oft eine besondere Kontextgebundenheit (vgl. DGfE 2017:

2), sodass die reine Darstellung via Metadaten des Öfteren nicht ausreicht und vielmehr eine umfangreichere Kontextualisierung notwendig ist.

2.4) Stellungnahmen der Fachgesellschaften

Die DFG hat mit den „Leitlinien zum Umgang mit Forschungsdaten“ im September 2015 die wissenschaftlichen Fachgesellschaften als Ansprechpartnerinnen beauftragt, disziplinspezifische Regelungen für die Handhabung von Forschungsdaten zu formulieren und deren Bereitstellung als Leistung anzuerkennen (vgl. DFG 2015: 2; siehe auch DFG 2009: 1). Fachgesellschaften, als verfasste Organisationen, repräsentieren oft einen wesentlichen Anteil der Forschenden eines Wissenschaftsfelds und machen sich für die jeweilige Fachdisziplin stark. Im Folgenden werden einige Verlautbarungen der Fachvertretungen von Disziplinen, die häufig mit sozialwissenschaftlichen qualitativen Daten arbeiten, dargestellt, um zu zeigen, wie sie sich zu jener Aufforderung verhalten haben.

Hierbei handelt es sich exemplarisch um die Deutsche Gesellschaft für Erziehungswissenschaft (DGfE), die Deutsche Gesellschaft für Psychologie (DGPs), die Deutsche Gesellschaft für Soziologie (DGS), die Deutsche Gesellschaft für Sozial- und Kulturanthropologie (DGSKA), die Deutsche Gesellschaft für Volkskunde (dgv) sowie das Fachkollegium 106 Sozial- und Kulturanthropologie, Außereuropäische Kulturen, Judaistik und Religionswissenschaft (Fachkollegium 106).

Die möglichen Mehrwerte in Form von dauerhafter Materialerhaltung, vertiefenden Analysen und einer vergrößerten Sichtbarkeit als auch die grundsätzliche Verantwortung, Interessierten die erhobenen Forschungsdaten zugänglich zu machen, werden von den Fachgesellschaften anerkannt (vgl. DGSKA 2019: 1).¹¹

¹¹ Die DGS nennt folgende Vorteile: „So fördern Sekundäranalysen die Transparenz des Forschungsprozesses, indem über Auswertungsverfahren und Interpretationen auf einer gemeinsamen Datengrundlage diskutiert und indem teilweise Metaanalysen oder Replikationen vorgenommen werden können, was über neue inhaltliche Erkenntnisse hinaus nicht zuletzt der Qualitätssicherung empirischer Forschung dienen kann. Nachnutzungen ermöglichen weiterhin die Analyse von (sozialen) Veränderungen bis hin zur Bearbeitung von Problemstellungen, die zum Erhebungszeitpunkt noch nicht absehbar waren bzw. nicht im Vordergrund standen. Auf diese Weise können ebenfalls Kooperationen zwischen Fächern bzw. Forschungsdisziplinen unterstützt werden. Zusätzliche Vorteile bestehen darin, dass eine intensive Nutzung bereits erhobener Daten nicht allein zeit- und kostensparend für die Forschenden ist, sondern auch auf der Seite der Forschungsteilnehmenden den Aufwand und die Belastungen reduziert. Schließlich können Sekundärdaten auch in der Lehre mit Gewinn eingesetzt werden.“ (DGS 2019: 3)

Für die Psychologie (DGPs) bestehen die Ziele laut Schönbrodt et al. in einer „Qualitätssicherung“, „Optimierung des Erkenntnisgewinns“ und „Maximierung des Kosten-Nutzen-Verhältnisses“ (Schönbrodt et al. 2017: 21).

Alle unterstützen die Bereitstellung von Forschungsdaten, wenn dieses auf Freiwilligkeit basiert. Anstelle von generischen Lösungen soll differenziert vorgegangen werden, die jeweiligen forschungsspezifischen Charakteristika Beachtung finden und Behinderungen des Datenerzeugungsprozesses vermieden werden (vgl. dgv 2018: 5; DGfE 2017: 3). Die Entscheidung über die Archivierung und Bereitstellung von Forschungsdaten und deren Eignung für eine sekundäre Nutzung sollte keinen Einfluss auf eine mögliche Förderung haben (vgl. DGfE 2017: 8). Es sollte (zeitliche) Flexibilität zum Beispiel im Hinblick auf Fristen der Bereitstellung ermöglicht werden (vgl. DGS 2019: 2, 9; DGSKA 2019: 2).

Die erfolgreiche Primärforschung als auch der Schutz vor Nachteilen aller am Forschungsprozess Beteiligten soll gegenüber den Interessen von Nach- und Weiternutzung unbedingt Vorrang haben (vgl. dgv 2018: 3; Fachkollegium 106 2019: 2). Die Heterogenität der Forschungsmethoden soll unhierarchisiert akzeptiert werden (vgl. DGSKA 2019: 2; DGfE 2017: 8). Der Vorrang der Primärforschung kann bedeuten, dass es in Einzelfällen möglich ist, dass bei ethischen oder rechtlichen Vorbehalten, z. B. Zustimmungsverweigerung für die Nachnutzung, auf die Datenarchivierung verzichtet bzw. die Zugänglichmachung der Daten jederzeit flexibel eingeschränkt wird und es zu einem nachträglichem Embargo oder gar zur Löschung der Daten kommen kann (vgl. dgv 2018: 4; Fachkollegium 106 2019: 2; Schönbrodt et al. 2017: 23 f.).

Der RatSWD unterscheidet drei Varianten der „nachhaltige(n) Sicherung von Forschungsdaten“: „ohne nutzungsfreundliche Dokumentation und ohne Nachnutzungsmöglichkeiten“, „mit nutzungsfreundlicher Dokumentation und mit restriktiven Nutzungsmöglichkeiten“ sowie „mit nutzungsfreundlicher Dokumentation für die Datennutzung außerhalb des Aufbewahrungsortes“ (RatSWD 2018b: 7).

Zugleich bedarf es einer individuellen Abwägung, ob sich der zusätzliche Aufbereitungsaufwand im Hinblick auf eine mögliche Nachnutzung lohnt, wobei generell eine erweiterte Ressourcenausstattung zum Zwecke des Forschungsdatenmanagements für Forscher_innen gefordert wird (vgl. DGS 2019: 5; DGSKA 2019: 3; dgv 2018: 5).

Die DGfE schlägt in diesem Zusammenhang drei Kriterien der Entscheidungsfindung vor: „die [...] Archivierungswürdigkeit der Daten unter inhaltlichen und methodischen Gesichtspunkten, die [...] Zulässigkeit ihrer Archivierung, bei der rechtliche und ethische Aspekte berücksichtigt werden, und die [...] Archivierungstauglichkeit, bei der die technische und informationelle Qualität der Daten“ (DGfE 2017: 6) zentral ist.

Von Seiten der Dateninfrastruktureinrichtungen muss zudem ein kostenloser und fortwährender Zugang zu den Daten gewährleistet werden (vgl. DGS 2019: 9). Des Weiteren weisen sie auf die Notwendigkeit von spezifischer Ausbildung und Beratung für Studierende und Forschende hin, was Fragen des Datenschutzes, Urheberrechts, Nutzungsrechts und die Entwicklung disziplinübergreifender Qualitätsstandards umfasst (vgl. DGfE 2017: 1 f.; DGS 2019: 6 ff.; dgv 2018: 5).

Zusammenfassend kann festgehalten werden, dass es sich bei diesen Äußerungen in der Hauptsache um fachpolitisch-strategische Papiere handelt, die Empfehlungen für eine Datenpraxis beinhalten, aber keine detaillierten Arbeitsprozesse beschreiben. Grundsätzlich stimmen die Fachgesellschaften der Datenarchivierung unter bestimmten Bedingungen zu, dies umfasst die Sicherstellung von Freiwilligkeit, keine Beeinträchtigung des Feldzugangs sowie eine ausreichende Ressourcenausstattung. Eine unveränderte Übernahme von Verfahren aus anderen (quantitativen) Fächern wird hingegen abgelehnt.

Neben diesen Überschneidungen existieren auch Diskrepanzen, die sich aus der Spezifik der jeweiligen Fächer ergeben. Die Stellungnahme der Deutschen Gesellschaft für Soziologie ist deutlich umfangreicher und an einigen Stellen konkreter, was darin begründet liegt, dass es in dieser Disziplin schon umfangreichere Erfahrungen mit dem Thema Forschungsdatenmanagement gibt. So gibt es explizite Bezugnahmen auf etablierte Anonymisierungsverfahren und Datenbereitstellungsprozesse, die im Bereich der quantitativen soziologischen Forschung schon lange Zeit existieren. Als ein Anwendungsfall werden zudem Replikationen auf Grundlage der archivierten Daten thematisiert (vgl. DGS 2019: 3 f.). Dies ist ein Ziel, das in den ethnologischen Fächern keine Rolle spielt.

Bei den ethnologischen Fächern ist die grundsätzliche Offenheit des Forschungsprozesses ebenso bedeutsam wie Positionen, die sich aus der Spezifik des Fachs ergeben z. B. bezüglich Auswahlkriterien und Nutzungs- sowie Verwertungsrechten (vgl. dgv 2018: 3 f.). Insbesondere die Debatten, inwiefern Wissen in der Forschung von den Beforschten koproduziert wird und was dies für Fragen des Eigentums in Hinblick auf Forschungsdaten bedeutet, sind weitestgehend nur für die ethnologischen Fachgesellschaften relevant. Eine frei zugängliche Veröffentlichung wird für den Großteil der Daten ausgeschlossen (vgl. DGSKA 2019: 1; dgv 2018: 5).

2.5) Etablierungsprobleme des Forschungsdatenmanagements und Kritik am Forschungsdatenparadigma

Ben Kaden nennt allgemeine Gründe, die die Akzeptanz des Forschungsdatenmanagements von Forschendenseite verhindern (vgl. Kaden 2018: 2-6): Diese reichen von einem generellen Desinteresse an der Thematik, über eine vermeintliche Beeinträchtigung der eigenen Wissenschaftsfreiheit, hin zu Sorgen vor einem Kontrollverlust. Daten gelten als Forschungskapital, weshalb sie einbehalten werden, um die exklusive Nutzung sicherzustellen (vgl. Huber 2019a: 14). Wissenschaftler_innen wollen vermeiden, dass ihre Daten für nicht intendierte Anliegen verwendet werden (vgl. Kaden 2018: 3).

Vielfach existiert eine Divergenz zwischen Aufwand und Nutzen einer Forschungsdatenpublikation, die mitunter hohen Arbeitsaufwand bedeutet, dem jedoch gefühlt kein Gewinn an wissenschaftlicher Reputation entgegensteht (vgl. Borgman 2010: 7; Kaden 2018: 6).

Des Weiteren spielt der Mangel an Ressourcen, institutioneller Unterstützung sowie der Mangel an geeigneter Beratung und Schulung, Infrastruktur sowie Standards eine Rolle (ebd.: 2, 4). Vorgaben werden als zu unspezifisch empfunden z. B. in Bezug auf den Datenschutz oder Nachnutzungsgrenzen (ebd.: 4 f.).

Schlussendlich können Forschungsthemen als zu sensibel gelten, um die Daten frei zugänglich zu machen, und es ist unklar, wer längerfristig verantwortlich für die entsprechenden Forschungsdaten ist (ebd.: 4).

Borgman nennt als dahingehende Bedenken, dass Daten falsch interpretiert, missbräuchlich verwendet oder ohne entsprechende Anerkennung angeeignet werden (vgl. Borgmann 2010: 9).

Rümpel erwähnt die „Unwissenheit über persistente und qualitative Verwaltung der Daten, Hemmnisse bezüglich der Datenspeicherung oder fehlende Transparenz der gespeicherten Daten in Repositorien.“ (Rümpel 2011: 25)

In Hinblick auf die Geschichtswissenschaften thematisiert Queckbörner den Zeitaufwand, die große Unsicherheit bezüglich Urheber- und Lizenzierungsfragen und die Sorge, dass die Herausgabe von Forschungsdaten vor Projektende zu einer Form des „Ideendiebstahls“ führen könnte (vgl. Queckbörner 2019: 55).

Neben diesen auf qualitative sozialwissenschaftliche Daten übertragbaren Gesichtspunkten gibt es auch noch einige spezifische Hindernisse.

Sterzer und Kretzer nennen: „Forschungsdaten und ihre Beziehungen zueinander (sollen) authentisch erhalten bleiben, der Daten- und Persönlichkeitsschutz involvierter Personen (muss) gewährleistet sein (und) die forschungsethischen Grundsätze [...] adäquat abgebildet werden.“ (Sterzer/Kretzer 2019: 112)

Eine zentrale Barriere bezüglich der Etablierung des Forschungsdatenmanagements ist der Mangel an Erfahrung. Bei vielen qualitativen Daten legen sich Forschende auf vielfältige Weise, samt ihren Unzulänglichkeiten, offen (vgl. Imeri 2018a: 75). So erwähnt Huber, dass Forschende häufiger die Angst geäußert hätten „im Archiv bloßgestellt zu werden.“ (Huber 2019a: 15) Die Ängste in Form einer „fear of exposure“ (TN1: 439) spricht die Sorge, dass beispielsweise Fehler in der Datenerhebung durch andere Forschende offensichtlich und kritisiert werden, wurde in einem Interview beschrieben. Die Konkurrenzsituation zwischen Wissenschaftler_innen führe dazu, dass es an einem „vertrauensvolle(n) Miteinander [...], wenn Probleme auftreten, [...] die wohlwollend miteinander zu diskutieren“ (ebd.: 442-43) mangle. Eine dahingehende fachinterne Debatte könnte Abhilfe schaffen.¹²

Erste praktische Erfahrungen würden zu einem Einstellungswechsel der Forschenden führen: „Die sagen dann das: Eigentlich ist das total gut und eigentlich hätten wir das schon immer machen sollen.“ (ebd.: 494-95) Eine angestrebte Normalisierung des Forschungsdatenmanagements würde zudem umfassen, dass es normal ist, vor der Forschung intensiv zu prüfen, welche Daten bereits erhoben worden sind, woran es oft mangeln würde. Ein „Diskurs [...] von Leuten, die Lust haben, Sekundärnutzung zu machen und das auch öffentlich zu diskutieren“ (ebd.: 564-66) könnte Erfahrungswerte generieren und darüber hinaus helfen, den von Opitz und Mauer in ihrer Machbarkeitsstudie als zentral beschriebenen Vorbehalt der Forschenden, „dass die eigene Forschungsfrage zu speziell wäre, als das sie sich vorstellen könnten, dass es dafür fremde Daten geben könnte“ (Opitz/Mauer 2005: 11), aufzulösen.

Zumindest kurz soll hier die wissenschaftliche Kritik an der Archivierung von qualitativen sozialwissenschaftlichen Forschungsdaten adressiert werden: Eine frühe Kritik formulierte der Soziologe Stefan Hirschauer, in welcher er eine Parallelisierung und kritiklose Adaption eines quantitativen Datenverständnisses für die qualitative Forschung problematisiert (vgl. Hirschauer 2014: 303 f.). Der wissenschaftliche Mehrwert dieser Daten sei, im Gegensatz

¹² Im Zuge des von mir besuchten Tagesworkshops „Oral-History Interviews als Forschungsdaten“ in Potsdam wurde in diesem Zusammenhang die Notwendigkeit der Etablierung einer „Ethik der Nachnutzung“ thematisiert.

zu quantitativen Daten, viel niedriger, wobei eine verpflichtende Archivierung das Vertrauensverhältnis zu den Beforschten beeinträchtigen würde: „Das Hauptproblem einer flächendeckenden Datenarchivierung ist also, dass die Archivierung der Daten ihre Produktion untergraben kann.“ (Hirschauer 2014: 310) Zwar erkennt er an, dass die Sicherung von Erinnerungsfragmenten sinnvoll ist, warnt jedoch vor einer sinnlosen Archivierung, wo Daten nur angehäuft werden, was mit einem enormen Maß an Ressourcenverschwendung einhergehe (ebd.: 301, 311).

Vier Gegenargumente sind für ihn zentral: Replikationsstudien, wie sie in Forschung mit hohem Standardisierungsgrad praktiziert werden, sind bei qualitativen Studien weder gewünscht noch möglich (vgl. Hirschauer 2014: 306). Überprüfbarer, transparenter und besser qualitätskontrolliert werden die Daten durch Archivierung ebenfalls nicht, die Datenmenge ist zudem vielfach zu umfangreich, als das Außenstehende sie überblicken könnten. Eine Sekundärnutzung wäre zudem sehr begrenzt, da ohne Kontextualisierung ein Sinnverlust droht. Eine umfangreiche Kontextualisierung hingegen würde personell sehr ressourcenintensiv sein. Die Wahrscheinlichkeit, dass Daten erneut genutzt werden ist kaum gegeben. Er sieht wenige Anwendungsfälle: Einsatz der Daten in der Lehre, Einzelfälle der Sekundärnutzung oder im Zusammenhang mit Mixed-Method-Forschungsprojekten (ebd.: 307 ff.).

Aus der Perspektive ethnologischer Forschung formuliert Pels eine Kritik, die hier jedoch aus Platzgründen nicht im Detail wiedergegeben werden kann. Er problematisiert die ethischen Implikationen der Datenkoproduktion dahingehend, wem Wissen in seinen verschiedenen Formen gehört, auf welcher (ethisch-rechtlichen) Grundlage und von wem Eigentumsentscheidungen zu treffen sind (vgl. 2018a: 394; 2018b: 411).¹³

Des Weiteren wird in Bezug auf die Methodologie vorgebracht, dass die Daten durch ihre Zuspitzung auf eine konkrete Forschungsfrage nicht darüber hinaus auswertbar sind und es Sekundärnutzenden an persönlicher Verbindung zum Studiengegenstand mangle (vgl. Opitz/Mauer 2005: 6 f.).

¹³ Zur Kritik an diesem idealisierten Kooperationsverständnis siehe Simpson/Sleeboom-Faulkner: „For instance, by contextualising coproduction in a more systematic way, we may be able to deal more transparently with research ‘collaborations’ that are not based on prima facie assumptions of equality, equal motivation and equal control over ensuing data. Indeed, an increasing proportion of ethnographic research cannot build on mutual trust [...] cannot be transparent and moreover, is based on negotiated, ambiguous ethical tenets, often motivated by incompatible interests. [...] A more holistic notion of the co-production of knowledge should also include the roles played by funders, universities and the regulation of research ethics and data protection.“ (2018: 402 f.)

Diese Kritik ist im Falle Hirschauers älter, mittlerweile haben die Forschungsdatenzentren Lösungen bzw. zumindest Ansätze zum Umgang mit den angesprochenen Problemen erarbeitet. Außerdem geht sie zu wenig darauf ein, was man mit der Sicherung von Daten auch Positives gewinnen könnte. Trotzdem berichtete eine der Befragten, dass weiterhin eine Skepsis von Fachseite wahrnehmbar sei. Forschende würden die Archivierung von Forschungsdaten mitunter immer noch als aufgezwungen wahrnehmen und die Potenziale in Zweifel ziehen. Insbesondere im Umgang mit den eigenen Materialien würden sie postulieren: „Das kann eh niemand verstehen“ (TN3: 121), weil es sich teilweise um einen zu persönlichen Blick auf die Gegenwart handele.

Trotz dieser Einschränkungen sind viele Forschende an der Archivierung und Nachnutzung von Forschungsdaten interessiert. Ergebnisse in diese Richtung liefert die von Imeri und Danciu durchgeführte Studie, wonach das Thema in der ethnologischen Fachcommunity bekannt ist und die Archivierung sowie Nachnutzung von Forschungsdaten positiv bewertet wird. Etwa 65 % der Befragten gaben an, dass sie sich eine Sekundärnutzung von Fremddaten vorstellen könnten und 70 % würden unter spezifischen Regelungen ihre Daten in Repositorien bereitstellen. Zentral ist die Gewährleistung des Persönlichkeits- und Datenschutzes. Die praktische Anwendung ist jedoch weniger weit verbreitet, ungefähr die Hälfte der Umfrageteilnehmer_innen hatte bislang keine Datenmanagementpläne erstellt (vgl. Imeri/Danciu 2017: 18, 22, 24). Ein ähnliches Bild zeichnet die bereits 2005 erstellte Machbarkeitsstudie von Opitz und Mauer, in welcher sie mit qualitativen Interviewdaten arbeitende Forscher_innen befragten: Von diesen gaben über 60 % der Befragten an, dass sie generell bereit wären bei beendeten bzw. laufenden Projekten ihre Daten für eine Nachnutzung zur Verfügung zu stellen, für kommende Projekte lag der Anteil bei über 80 % (vgl. Opitz/Mauer 2005: 15).

Betrachtet man jedoch den Status Quo, ist mitunter eine Differenz zwischen dem, was vonseiten der Forschungsförderung (zukünftig) gewünscht ist, was an Infrastruktureinrichtungen erörtert wird, den vorhandenen technischen Möglichkeiten und dem aktuellen methodischen Vorgehen und der Fachkultur der einzelnen Wissenschaftsdisziplinen zu konstatieren (vgl. Harbeck et al. 2018: 125). So resümiert der RatSWD in einer Veranstaltungseinladung für den Workshop „Archivierung und Zugang zu qualitativen Daten“: „Die institutionelle Struktur für die Archivierung und Verfügbarmachung qualitativer Forschungsdaten ist in Deutschland immer noch ein Desiderat. Speziell für qualitatives Material bestehen bislang wenige Möglichkeiten, dieses nach Projektabschluss langfristig und nachhaltig zu archivieren.“ (RatSWD 2018a: 1) Auch

Imeri beschreibt ein „einigermaßen ernüchterndes Bild“ (Imeri 2018b: 218). Es mangle an adäquaten Standards und Repositorien sowie auf Seite der Forschenden oft an angemessenen Speicherroutinen. Des Öfteren verbleiben die Daten bei Forschenden oder werden nur mittels persönlicher Kontakte weitergegeben, sodass das Nachnutzungspotenzial nicht ausgeschöpft werden kann (ebd.). Trotz dieser Einschränkungen gibt es zahlreiche Infrastruktureinrichtungen, die Lösungsansätze für den Umgang mit qualitativen sozialwissenschaftlichen Forschungsdaten entwickelt haben bzw. entwickeln, die im Folgenden näher untersucht werden.

3.) Methodisches Vorgehen

3.1) Material und Methodik (Forschungsdesign und Operationalisierung der Fragestellung)

Wie bereits beschrieben beschäftigt sich die vorliegende Arbeit mit den Spezifika im Umgang mit qualitativen sozialwissenschaftlichen Daten im Kontext der Archivierung und Nachnutzung, welche Lösungsansätze entwickelt worden sind und wie mit vorhandenen Übertragungs- und Anwendungsproblemen umgegangen wird. Die vorliegende Untersuchung basiert auf qualitativen Interviews, die darauf abzielten, dass Expert_innen umfangreich und trotzdem fokussiert über ihre Einrichtungen und ihre persönliche Sichtweise auf das Thema Forschungsdatenmanagement berichten.

Aus dem vorher gesichteten Material wurden Themenschwerpunkte gebildet, die Operationalisierung erfolgte, indem diese in eine Reihe von spezifischen Interviewfragen übersetzt wurden, die an den Erfahrungsschatz der Interviewten anknüpften. Diese Fragen lassen sich neben rahmenden Fragen in zwei größere Blöcke unterteilen. Der erste Block umfasste die Themenbereiche Sammeln, Archivieren, Finden, Bereitstellen und der zweite Block bezog sich auf den Komplex der Nachnutzung. Die Interviews beinhalteten Elemente eines systematisierenden und explorativen Expert_inneninterviews, das auf die Vermittlung von exklusivem, aus der Praxis gewonnenen Erfahrungswissen abzielte. Zugleich umfassten sie aber auch Bestandteile eines theoriegenerierenden Interviews, das sich „auf die Erhebung von „Deutungswissen“ also jenen subjektiven Relevanzen, Regeln, Sichtweisen und Interpretationen des Experten“ (Bogner/Menz 2005: 43 f.) bezog. Die Methode war ideal geeignet, um weiteres Kontextwissen zu generieren und gleichzeitig in einem offenen dialogischen Prozess umfassende Einschätzungen zu erfragen, die sich durch standardisierte Befragungen wesentlich unzureichender eruieren ließen (vgl. Hopf 2012: 350).¹⁴

¹⁴ „Das Prinzip der Offenheit verlangt hier zunächst die Auswahl von Methoden, die den Akteuren im Forschungsfeld genügend Raum lassen, ihre Perspektiven in das Forschungsprojekt einzubringen.“ (Rinsdorf 2013: 71)

3.1.1) Das Expert_inneninterview: Expert_innenwissen und Expert_innenauswahl

Die Identifizierung relevanter Expert_innen, die Entwicklung des Leitfadens sowie ein parallel stattfindender Qualifizierungsprozess als Interviewer erfolgte systematisch und theoriegeleitet. Er basierte auf einer von Pfadenhauer geforderten „relativ detaillierten ethnographischen(n), Inventarisierung des Forschungsfeldes“ (Pfadenhauer 2005: 125). Dies geschah mittels der im Forschungsprozess gesichteten (Fach-)Literatur, die Auskunft darüber gab, welche relevanten Organisationen und Akteur_innen es in diesem Feld gibt sowie durch handschriftlich notierte Vorabgespräche mit verschiedenen Mitarbeiter_innen des FID Kultur- und Sozialanthropologie, die das Thema Forschungsdatenmanagement in den ethnologischen Fächern intensiv bearbeiten. Daneben besuchte ich den Tagesworkshop „Oral-History Interviews als Forschungsdaten“ des Leibniz-Zentrums für Zeithistorische Forschung in Potsdam, der sich u. a. mit der Erschließung, Kontextualisierung und Zugänglichmachung von Forschungsdaten und deren potenzieller Nachnutzung beschäftigte. Weitere Informationen und tiefergehende Einblicke konnten in einem vom 9.-13.3.2020 speziell organisierten Erasmus+Stafftraining an der UB Wien bzw. am Institut für Kultur- und Sozialanthropologie gewonnen werden.

Schlussendlich konnten auf dieser Grundlage sieben Interviews realisiert werden, die ausgewählt wurden, um die Vielfalt im Bereich der qualitativen Forschungsdaten abzudecken. Es handelte sich um Personen, die ihr privilegiertes Feldwissen aufgrund ihrer Involviertheit und ihres „relationalen Status“ (Meuser/Nagel 1991: 443) in Form von zentralen Funktionen im Feld des Forschungsdatenmanagements erworben haben (vgl. Meuser/Nagel 2009: 468 f.; Werner 2013: 128, 143). Sie konnten Auskunft über betriebs- und kontextrelevante Prozesse sowie deren Implementierung geben und die Praxis des Feldes aktiv gestalten (vgl. Bogner/Menz 2005: 46).

Fünf der Interviewten arbeiteten dabei in Einrichtungen, die sich mit der Archivierung eines breiten Spektrums qualitativer Daten, hauptsächlich aus den Sozialwissenschaften (den Ethnologien, Politikwissenschaft, Soziologie) und der Geschichtswissenschaft, beschäftigen, wobei die jeweilige Fachspezifik immer wieder andere Lösungsansätze bedingte. Zwei der Interviewten arbeiteten eher im Bereich der quantitativen Forschungsdaten, wobei es in ihrer sozialwissenschaftlichen Institution einen kleineren Anteil an qualitativen Daten gab. Dieser etwas abweichende Hintergrund stellte sich als sehr

erkenntnisfördernd heraus, bot er zum einen eine Vergleichsfolie und dadurch, dass das Forschungsdatenmanagement in diesem Bereich länger etabliert ist, Einblicke dahingehend, welche Verfahren in Zukunft auch für die qualitativen Daten von Belang sein könnten. Die Spannweite der befragten Einrichtungen reichte von kleinen an Universitäten angesiedelten Institutionen, die gerade erst in der Implementierung begriffen sind, über etablierte Institutionen, die schon über langjährige Erfahrungswerte verfügen, bis hin zu größeren sozialwissenschaftlichen Infrastruktureinrichtungen mit zahlreichen Mitarbeiter_innen.

3.2) Vorgehensweise Expert_inneninterviews

3.2.1) Feldzugang und Interviewdurchführung

Die primäre Kontaktaufnahme erfolgte via E-Mail, wobei knapp die Forschungsmotivation und die Wichtigkeit der Interviewteilnehmer_innen angeschnitten wurden (vgl. Bogner et al. 2014: 39). Insgesamt wurden sieben Interviews, vier vor Ort und drei via Telefon mit einer Dauer zwischen ca. 45-120 Minuten geführt. Die meisten Interviews umfassten einen Zeitraum von knapp 45 Minuten bis maximal 75 Minuten, wobei die zeitliche Diskrepanz insbesondere durch den Grad der Detailliertheit der Beschreibungen von Seiten der Befragten begründet war. Das zweistündige Interview entwickelte sich spontan aus einem Gespräch heraus und wurde deshalb handschriftlich mitprotokolliert. Es umfasste, trotz seiner freieren Form, die gleichen Themenkomplexe, die auch bei den anderen Interviews abgefragt wurden. Der ursprüngliche Plan sah vor, dass alle Interviews in Form von persönlichen Gesprächen in den jeweiligen Einrichtungen stattfinden sollten, was sich jedoch aufgrund der Corona-Pandemie leider nicht realisieren ließ. Die telefonischen Interviews hatten den Nachteil, dass wichtige Elemente des Interviews wie Gestik und Mimik nicht beobachtbar waren und beispielsweise schwieriger zu bestimmen war, wann die Interviewten fertig geantwortet hatten.

Innerhalb der Interviewsituation dienten zentrale Forschungsmethodenkriterien als Orientierung (vgl. Schmidt-Lauber 2007: 172): Das Kommunikationskriterium sah vor, dass den Expert_innen ausreichend Platz eingeräumt sowie eine empathische und geduldige Gesprächsführung angestrebt wurde. Das Prinzip der Offenheit beinhaltete, ohne fest gefügte Vorstellungen und mit der Bereitschaft für einen Wechsel des Blickwinkels die

Interviews zu bestreiten. In einer Einleitungsphase wurde die Zustimmung zur Interviewaufzeichnung mit Hilfe eines akustischen Aufnahmegeräts eingeholt, das sich aufgrund der großen Datenmenge auch insgesamt als adäquates Aufzeichnungsmittel herausstellte. Darüber hinaus wurde jeweils der Interview-/Forschungsablauf erläutert und erste Rückfragen vonseiten der Expert_innen, beispielsweise in Hinblick auf mögliche Anonymisierungen oder den Studienhintergrund, beantwortet. Den Interviewten wurde außerdem zugesagt, dass sie die abgeschlossene Studie via E-Mail zugesandt bekommen.

3.2.2) Erhebungsinstrument Leitfaden

Die Interviews basierten auf einem Leitfaden, der half, das Themenfeld zu strukturieren und zudem als praktisches Hilfsmittel während der Befragungssituationen fungierte (vgl. Bogner et al. 2014: 27). Diese Interviewform wurde gewählt, weil sie am besten dazu geeignet schien, innerhalb der beschränkten Interviewzeit unterschiedliche Themen mit Sicherheit anzusprechen (vgl. Gläser/Laudel 2010: 111, 187). Zudem konnte das Interview so auf das tatsächliche Erkenntnisinteresse fokussiert bleiben und nicht in wenig relevante Bereiche abdriften. Innerhalb des Interviews wurde darauf geachtet, eine Balance zwischen dem Ziel der Offenheit (der Fragen) und der Strukturierung mittels des Leitfadens herzustellen. Damit wurde einerseits die Möglichkeit geschaffen, sich in beschränkter Zeit über unterschiedliche Themen gezielt zu informieren, gleichzeitig, aber auch offen für neue Impulse zu sein (vgl. Gläser/Laudel 2010: 131). Auf diese Weise war zudem eine bessere Vergleichbarkeit mit den anderen Interviews gegeben. Die Erstellung eines Leitfadens setzt „umfassendes und einschlägiges Wissen“ (Pfadenhauer 2005: 121) bei dem Interviewer als „Quasi-Experten“ (ebd.) voraus, weshalb eine ausführliche Beschäftigung mit dem Gegenstand stattgefunden hat, um so ein Gespräch „auf gleicher Augenhöhe“ (ebd.: 127) führen zu können.

Mit diesem Leitfaden wurde flexibel umgegangen, sprachliche Formulierungen, Fragereihenfolge und die sich ergebenden Nachfragen vonseiten des Forschers wurden spezifisch an die Expert_innen und den Gesprächsverlauf angepasst (vgl. Hopf 2012: 351; Werner 2013: 130). Auf diese Weise und durch einen Pretest wurde eine „Leitfadenbürokratie“ (Hopf 2012: 358) verhindert, sprich ein zu langer Leitfaden sowie das fest gefügte und wenig tiefgreifende Abfragen der Punkte. Eine individuelle Anpassung basierte u. a. auf der Sichtung von Publikationen der Expert_innen bzw. der Institution, woraus zentrale Zitate als Gedächtnisstütze in den Leitfaden integriert wurden. Der

Leitfaden orientierte sich dabei an den zuvor aufgestellten Forschungsfragen und zielte auf deren umfassende Beantwortung ab. Der Interviewprozess war insofern iterativ, als dass sich Interviewphasen und die interpretative Auswertung abgewechselt haben, die dann mitunter ergänzende Detailfragen für den Leitfaden geliefert hat. Grundsätzlich gab es jedoch keine größeren Abweichungen vom Leitfaden und den darin integrierten übergeordneten Fragekomplexen.

Die Strukturierung der Interviewfragen begann jeweils mit einer einfachen Einstiegsfrage über das Betätigungsfeld, die den Interviewten die Möglichkeit geben sollte, sich „warm zu reden“ (Bogner et al. 2014: 61). Jeder Themenabschnitt fing mit einer offen gehaltenen Frage an, die als Erzählstimulus fungierte, um den Expert_innen zu ermöglichen eine assoziative Antwort bezüglich ihrer Sichtweise zu geben. Dabei war die Wahrung von Neutralität, die Vermeidung von Suggestivfragen sowie eines zu dominanten Kommunikationsstils zentral (vgl. Gläser/Laudel 2010: 135). Entweder schlossen sich dann im Anschluss, zwecks Vervollständigung oder Stellungnahmen, sondierende Nachfragen an oder die Unterhaltung wurde wieder in Richtung des Interviewleitfadens gelenkt. Waren Dinge unklar und bedurften weiterer Interpretation diente das „Spiegeln von Aussagen“ (Werner 2013: 133) in Form von ersten Deutungen und Zusammenfassungen als ein Mittel, um ein tiefgreifenderes Verstehen der Expert_innenaussagen zu erreichen. Trotz des Leitfadens wurde darauf geachtet, nicht in starre Fragemuster zu verfallen, sondern vielmehr situationsangemessen mit diesem umzugehen (vgl. Schmidt-Lauber 2007: 177). Die Stoßrichtung der Interviews bewegte sich in Form eines „Fragetrichters“ (Werner 2013: 132) von eher offenen erzählregenden Fragen hin zu fokussierenden eher geschlossenen Detailfragen. Erstere zielten auf umfassende (bewertende) Aussagen und Charakterisierungen, zweitere auf kürzere Antworten ab. Jedes Interview endete mit der sich als sehr produktiv herausstellenden bilanzierenden Abschlussfrage, ob die Interviewten den Eindruck hätten, ob noch etwas Wichtiges im Interview nicht ausreichend erwähnt worden sei (vgl. Bogner et al. 2014: 61). Hierbei stellte sich heraus, dass der Leitfaden die relevanten Fragen abfragen konnte und den Befragten keine wesentlichen Aspekte fehlten, vielmehr nutzten sie die Gelegenheit, um Sachverhalte hervorzuheben, die ihrer Meinung nach besondere Aufmerksamkeit verdienten. Dies umfasste Fragen der Ausbildung, der dauerhaften Finanzierung oder einer kriterienbasierten Datenauswahl im Archivierungsprozess. Manchmal entwickelten sich nach dem eigentlichen Interview noch Gespräche, die in kurzen Memos ebenfalls aufgezeichnet wurden.

3.3) Auswertung

3.3.1) Transkription

Die ersten Transkriptionen begannen bereits zu Beginn des Untersuchungsprozesses, um so zeitnah frühe Interpretationen zu generieren, auf Leerstellen zu achten und schließlich den eigenen Interviewstil einer Überprüfung zu unterziehen (vgl. Schmidt-Lauber 2007: 182). Die angewandten Transkriptionsregeln und der Umfang der Transkription wurden, nach dem Führen der ersten Interviews, festlegt und basierten auf Überlegungen, welche Informationen zur Bearbeitung des Forschungsvorhabens nötig waren (vgl. Werner 2013: 135).

Die Transkription des Interviewmaterials erfolgte vollständig, um einerseits den interpretativen Vorgriff auf den Text durch Teiltranskriptionen zu vermeiden und weil nach dem ersten erneuten Hören der Eindruck entstand, dass fast alle Teile für die Forschung relevante Hinweise enthielten (vgl. Bogner et al. 2014: 80; Meuser/Nagel 1991: 456).

Die spezifische Transkription sollte „die Analyse von Themen und Inhalten der aufgezeichneten Vorgänge ermöglichen, nicht aber von Interviewereinflüssen, Betonungen.“ (Knoblauch 2012: 159) Die Konzentration auf diese verbalen Inhalte führte zur Entscheidung, auf die Transkription von prosodischen Interviewbestandteilen wie Lautstärke und Tonhöhe, parasprachlichen Merkmalen beispielsweise Räuspern oder Lachen sowie die Darstellung außersprachlicher Komponenten wie Mimik und Gestik zu verzichten, da diese nicht Teil des Analysefokus darstellten (vgl. Mäuser/Nagel 1991: 455; Kowal/O'Connell 2005: 444). Die Reduktion der Datenmenge im Übersetzungsprozess von Tonaufnahme hin zum geschriebenen Transkript sollte zugleich eine bessere Les- und Interpretierbarkeit für Außenstehende ermöglichen. Umgangssprachliche Elemente wie Elisionen und Assimilationen wurden behutsam an die Standardorthografie angepasst (vgl. Kowal/O'Connell 2005: 441).

Vor den Interviews wurde von den Interviewten mehrfach darauf hingewiesen, dass es sich bei den von ihnen gemachten Aussagen über die eigene Arbeit und Organisation um sensible Themen handelt. Dies umfasste beispielsweise die Entwicklung neuer Angebote, im Zuge dessen auftretende Probleme und Fragen der Finanzierung. Zum Schutz vor Nachteilen und zur Sicherung des Vertrauens werden aus diesem Grund die Aussagen der Interviewten im Folgenden anonym mit TN1 etc. wiedergegeben und eine Pseudonymisierung der

Transkripte durchgeführt (vgl. Medjedović/Witzel 2010: 76 ff.). Ziel war die Reduktion und das systematische Verändern von Informationen, die Rückschlüsse auf die Interviewten ermöglichen könnten (vgl. Gebel et al. 2015: 13). Hierzu wurde ein Pseudonymisierungsschlüssel entwickelt: So wurden Namen von Beteiligten, Ortsangaben und die Namen von Infrastruktureinrichtungen durch Platzhalter ersetzt, um Rückschlüsse auf die Identität der Befragten zu verhindern. Dies besaß den Vorteil, dass über verschiedene Datensätze verteilte Bezüge erhalten blieben. An wenigen Stellen wurde zudem mit Vergrößerungen gearbeitet und sehr ausführliche Beschreibungen von spezifischen Organisationen vermieden, dabei jedoch darauf geachtet, dass die Eigenschaften der Einrichtung kenntlich blieben.

3.3.2) Codierung

Im Anschluss an die durchgeführten Transkriptionen folgte ein an Schmidt (2012) und Bogner et al. (2014) orientierter iterativer Codierungsprozess auf Basis der mehrmaligen Materiallektüre. Die mit MAXQDA erfolgte Auswertung basierte auf der Bildung thematischer Blöcke, die über die gesamten Transkripte verstreut waren, nicht auf der Abfolge der Aussagen innerhalb der Interviews (vgl. Meuser/Nagel 1991: 453). Ziel war es, auf der Grundlage der vorliegenden Interviews gemeinsame Interpretationsmuster und Wissensstände, aber auch spezifische Problemlösungsansätze herauszuarbeiten, die der gezielten Beantwortung der Forschungsfragen dienen (vgl. Bogner et al. 2014: 78).

In einem ersten offenen Codierungsschritt wurde das Material aufgebrochen, systematisch durchgesehen und mit frei gewählten und assoziativen Stichworten, die mitunter die Begrifflichkeiten der Befragten übernahmen, versehen. Zudem wurden wesentliche Aussagen paraphrasiert oder durch Kurzmemos ergänzt. Die thematische Auswertung wurde individuell für jedes Interview durchgeführt, sie wurden an dieser Stelle noch nicht verglichen und zusammengefasst. Ziel war das Erkennen und Benennen erster Muster und zentraler Themen (vgl. Schmidt 2012: 449). Dabei wurde vermieden „das Material [...] auf die eigenen theoretischen Vorannahmen zuzuschneiden. Indem die Auswertung darauf reduziert wird, nur nach solchen Textstellen zu suchen, die sich als Beleg oder Illustration für die eigenen Vorannahmen eignen.“ (Schmidt 2012: 450) Ziel war es, trotz der vorstrukturierenden Leitfragen für neuartige und überraschende Aspekte aufgeschlossen zu sein.

Im Anschluss wurden auf Basis der entwickelten Codes gemeinsame abstraktere Kategorienbezeichnungen entwickelt, worunter einzelne Codes sortiert wurden (vgl. Bogner et al. 2014: 79). Dies ermöglichte eine Bündelung von themengleichen Passagen, deren Beziehung untereinander sowie eine Übersicht bezüglich des Spektrums verschiedener Kategorienausprägungen (vgl. Schmidt 2012: 451; Schmidt-Lauber 2007: 184). Durch das thematische Ordnen wurden über die Interviews verstreute, aber zusammenhängende Information nicht nur ausdifferenziert und integrierend verdichtet, sondern gleichzeitig zentrale Belegstellen herausgearbeitet (vgl. Bogner et al. 2014: 74).

In einem dritten Schritt wurden auf Grundlage dieser abstrakten Bezeichnungen die Interviewtranskripte erneut betrachtet und offene Interviewpassagen passenden Kategorien zugeordnet (vgl. Schmidt 2012: 452). Diese umfassenden Kategorien wurden schließlich systematisiert und bezüglich ihres größeren Zusammenhangs, Deutungs- und Relevanzmustern interpretiert und verglichen (vgl. Bogner et al. 2014: 79 f.). Die entwickelten Hauptkategorien entsprechen dabei den im kommenden Abschnitt folgenden Unterkapiteln und strukturieren somit die Analyse der Forschungsergebnisse.

4.) Analyse der Forschungsergebnisse

Der folgende Teil beschreibt, anhand der durch Interviews und Forschungsliteratur gewonnenen Informationen, relevante Themen des Managements qualitativer sozialwissenschaftlicher Forschungsdaten. Er zeigt, wie Infrastruktureinrichtungen mit den damit einhergehenden Erfordernissen umgehen und welche Lösungsansätze sie bei der Archivierung sowie der Nachnutzung der Daten entwickelt haben. Dies umfasst eine (forschungs-)ethische, methodische, rechtliche sowie praktisch-technische Ebene. Konkret dargestellt ist, wie mit dem heterogenen Material bezüglich der notwendigen Kontextualisierung, der Klärung von Persönlichkeitsschutz, Urheberrecht sowie der Auswahl der Daten zur Archivierung umgegangen wird und welche diesbezüglichen Nachnutzungsszenarien sowie Hindernisse existieren.

4.1) Fragen (forschungs-)ethischer und methodischer Natur

Rechtliche und (forschungs-)ethische Aspekte sind, als sich überschneidende Bereiche, gleichermaßen für das Forschungsdatenmanagement qualitativer sozialwissenschaftlicher Daten relevant. Eine zentrale forschungsethische Maxime ist „Schaden zu vermeiden und die Rechte und Anliegen der Teilnehmer/innen zu achten.“ (Corti 2018: 22) Der Quellenschutz ist ein wichtiges Element qualitativer Forschung, der nicht auf formalrechtliche Kategorien heruntergebrochen werden kann. In Bezug auf das Forschungsdatenmanagement sind forschungsethische Aspekte bei der Erhebung, der Auswertung und der Veröffentlichung der Daten permanent von Relevanz. Folgende Momente müssen dabei reflektiert werden:

Die bereits beschriebene Her- und Sicherstellung einer Vertrauensbeziehung zwischen Forschenden und Beforschten ist nicht nur ethisch geboten, sondern auch unerlässlich, um die erfolgreiche Durchführung der Forschung überhaupt erst zu gewährleisten. Die Infrastrukturen müssen diese fachwissenschaftlichen Positionen reflektieren und darauf in Form von Angeboten reagieren, damit Datengeber_innen bereit sind, ihre Forschungsdaten zu übergeben.

Zudem ist es, insbesondere aus ethnologischer Sicht, in vielen Fällen notwendig, über die Machtstrukturen, in die die Forschung(-sdaten) zwangsläufig eingebettet sind, nachzudenken und eine respektvolle Behandlung auf gleicher Augenhöhe mit den untersuchten Personen zu gewährleisten und diese auch in das Forschungsdatenmanagement selbst zu integrieren. Dies kann konkrete Forderungen nach Beteiligung bei der Anonymisierung und der Datenzugänglichkeit bis hin zur Löschung von Datensätzen umfassen. Dazu gehört bestenfalls auch die Über- und Rückgabe von Daten und Veröffentlichungen und die Möglichkeit Forschungsergebnisse kritisch zu kommentieren (vgl. Eberhard/Kraus 2018: 47 f.). Dies soll, im Einzelfall und je nach Forschungsfeld, Fehlannahmen korrigieren und so die Forschungsqualität verbessern.

Für Forscher_innen als Verwalter_innen potenziell gefährdender Daten ist die kritische Prüfung möglicher Risiken und deren Reduzierung zentral (vgl. Sterzer/Kretzer 2019: 112). Im Kontext des Umgangs mit Forschungsdaten bedeutet dies, dass Schadensvermeidung, via datenschutzrechtlichen Fragestellungen, eine zentrale Stelle einnimmt. Insbesondere die Weitergabe von Daten zum Zwecke einer möglichen Nachnutzung muss ebenfalls unter forschungsethischen Gesichtspunkten betrachtet werden. Daten sind für die Teilnehmer_innen potenziell riskant, sei es durch das soziale Umfeld oder staatlicherseits, auch wenn sie sich nicht direkt auf Einzelpersonen beziehen, zum Beispiel im Rahmen von Aktivitäten am Rande der Legalität sowie in Politik, Migration oder Gesundheit betreffenden Zusammenhängen (vgl. Imeri 2018b: 221). Wenn bestimmte Daten bekannt werden, kann dies weitreichende negative Folgen haben, wie der Verlust des Arbeitsplatzes oder juristische Konsequenzen beispielsweise in Bezug auf das Straf-, Zivil- oder Aufenthaltsrecht (vgl. von Unger 2018: 94). Eberhard und Kraus (2018: 48) weisen zudem darauf hin, dass die Sichtbarkeit von Daten bezüglich Regionen wie Afghanistan oder dem Irak dazu führen kann, dass Kulturgüter lokalisiert und zerstört werden können oder Informant_innen um ihr Leben fürchten müssen. Eine Person erwähnte im Zuge des Erasmus+-Programms, dass sie in Bezug auf Material aus Afghanistan, um Zerstörung zu vermeiden, bestimmtes Material nicht veröffentlicht bzw. auf die Beschreibung durch GPS-Daten verzichtet. Was als sensibles und potenziell schädigendes Datum gilt, ist permanenten Wandlungen unterworfen, die zum Zeitpunkt der Datenübergabe kaum absehbar sind, sodass diese Gefahr immer mitbedacht werden muss (vgl. Harbeck et al. 2018: 132). Von Unger weist zudem darauf hin, dass zukünftige technische Entwicklungen noch nicht absehbar sind. Die Vernetzung der Digitaltechnik könnte beispielsweise dazu führen, dass die

Anonymisierung visueller Daten durch Software zur Gesichtserkennung und Ortsbestimmung konterkariert werden könnte (vgl. von Unger 2018: 94). Eine beschränkungsfreie Zugänglichkeit der Daten ist deshalb nicht realisierbar (vgl. Eberhard/Kraus 2018: 46). Vielmehr sind, nach einer Risikoabwägung, Zugangsbedingungen notwendig.

Auf methodischer Ebene müssen auch die etwaigen Rückwirkungen des Datenmanagements auf die Forschungspraxis diskutiert werden. So weist Imeri auf eine mögliche „Schere im Kopf“ (Imeri 2019: 49) hin, die entstehen könnte, wenn die Forschenden im Hinterkopf haben, dass ihre Daten möglicherweise von anderen betrachtet werden. Dies könnte die Art und Weise, wie innerhalb von Interviews agiert wird oder wie Dinge beschrieben und aufgezeichnet werden verzerren (vgl. Imeri 2017: 173). Innerhalb einer Einrichtung wird in diesem Zusammenhang deshalb eine explorative Anwendungsstudie durchgeführt, um herauszufinden, wie die Datenarchivierung und potenzielle Nutzung durch Dritte auf die Forschung, beispielsweise bezüglich Teilnahme- und Antwortverhalten der Beforschten, zurückwirkt (vgl. TN2: 624-32).

Neben diesen Aspekten, die eher eine Beschränkung der Archivierung und Nachnutzung nahelegen, existieren jedoch auch Argumente, die dafürsprechen, dass es eine ethische Verpflichtung zur koordinierten Archivierung und Zugänglichmachung gibt. Dies sehen auch zuständige Fachgesellschaften so: „Gleichzeitig besteht eine generelle Verantwortung des Fachs, Forschungsmaterial interessierten NutzerInnen zugänglich zu machen, sofern die institutionellen und technischen sowie ethischen Voraussetzungen dafür gegeben sind.“ (DGSKA 2019: 1)¹⁵ Durch die Sekundärnutzung ist die Vermeidung von nochmaligen psychischen sowie physischen Belastungen durch erneute Beforschung möglich. Smioski nennt als Beispiel biografische Interviews mit älteren Zeitzeug_innen in einer ländlichen Region Österreichs, wo diese über Kriegs- und Nachkriegserfahrungen berichteten, die sonst verloren, gegangen wären (vgl. Smioski 2011: 221 f.).

Des Weiteren umfasst die Archivierung die nachhaltige Bewahrung von kulturellem Erbe, Zeitdokumenten sowie Wissen beispielsweise im Falle politischer Veränderungen. Ein Beispiel ist die umfassende Materialsammlung des Forschers Werner Finke, der über drei

¹⁵ Vgl. auch EASA (2018: 1) und die CARE Principles for Indigenous Data Governance (RDAIIDSIG 2019).

Jahrzehnte in die kurdischen Gebiete in der Türkei gereist ist und dort intensive Forschung in dieser stark von politisch-militärischen Interventionen gezeichneten Region betrieben hat. Nun ist vom Institut für Sozialanthropologie (ISA) der Österreichischen Akademie der Wissenschaften (ÖAW) ein digitales Archiv samt freiem Onlinezugang in Planung, dessen Ziel ist verschiedene Materialsammlungen zusammenzuführen, „um das kulturelle Erbe zu bewahren“, da jene einen „außerordentlichen Wert für die betreffenden Regionen“ haben (ÖAW ISA 2020). „Damit soll den Menschen aus den Regionen, in denen Werner Finke gearbeitet hat, diese umfassende Dokumentation ihrer Kultur wieder «zurückgegeben» werden.“ (Six-Hohenbalken/Emir 2016)

Ein weiteres Beispiel stellen die National Anthropological Archives dar, wo Feldmaterialien zunehmend nicht nur von Wissenschaftler_innen konsultiert werden, sondern insbesondere von Native Americans, die ihr eigenes kulturelles Erbe untersuchen. Die Daten für die genealogische Forschung, für die Unterstützung bei der Wiederbelebung gefährdeter Sprachen und für die Suche nach einer historischen Grundlage für Land- oder andere Besitzansprüche, nutzen (vgl. Leopold 2008: 5). Diesem Interesse kommt das Archiv mit der aktiven Förderung von sogenannten Wissensrückführungsprogrammen (knowledge repatriation programs) entgegen (vgl. Leopold 2008: 9).

Vielfach sind Daten nur unzureichend bei Forschenden gesichert und drohen durch ein Projektende, Verrentung oder gar den Tod der Forschenden nicht mehr zugänglich zu sein oder gänzlich zu verschwinden (vgl. Imeri 2017: 168). Hier zeigt sich, dass es weniger darum geht, ob die Daten überhaupt archiviert werden sollen, sondern vielmehr darum unter welchen Bedingungen und Schutzmaßnahmen sie gesichert und zur Nachnutzung bereitgestellt werden sollen.

Eine Interviewte konstatierte, dass die Archivierung durch zentrale Einrichtungen sinnvoll sei, da in Privatarchive und in zeitlich begrenzten Forschungsprojekten, die Chancen für eine angemessene Nachnutzung mangelhaft sind (vgl. TN3: 3-5). Bestenfalls gingen die Daten bislang in Institutionen wie Archive oder Bibliotheken aber auch private Sammlungen über, wo sie aber oft nicht in einem ausreichenden Maße erschlossen und durchsuchbar waren. Aufwand, Fachwissen und ausreichend zeitliche Ressourcen waren zur Nutzung nötig, sodass „der Zugang zu Quellen und Daten ein Stück weit esoterisch und elitär“ blieb (Eberhard/Kraus 2018: 42). Der historische Wert der Materialien erfordert, dass es möglich ist, mit Distanz erneut auf die Daten zu blicken. Insbesondere auch, weil ethnografisches Material, in den Ethnologien, kollaborativ mit und für die beforschten

Gruppen erstellt wird. Bislang wurde jedoch kein kollaborativ erzeugtes Material, das von der Einrichtung archiviert wurde, zurück in das Feld, wo es entstanden ist, mitgenommen, um dort eine gemeinsame (Re-)Analyse durchzuführen (vgl. TN3: 66-74).

Eine andere interviewte Person berichtete, dass ein Ziel darin besteht, dass zwischen Forschenden und Beforschten ein Kontakt hergestellt wird und eine Zugänglichmachung des Materials anvisiert wird. Dies erfolgt bereits indem Wissenschaftler_innen Links zu Forschungsdaten an die Beforschten bzw. zumindest an Forschungscommunities, die ebenfalls zu dem Forschungsgegenstand arbeiten, weitergeben (vgl. TN4: 264-70). Die Einrichtung sieht die Forschenden hierbei ethisch verpflichtet diese Weitergabe zu leisten, „das sollte eigentlich ein No-Go sein, das nicht zu machen.“ (TN4: 278) Werden Lösungen erarbeitet, die den Forschungssubjekten Datenzugang oder gar ein Mitspracherecht ermöglichen, bietet dies darüber hinaus Ansatzpunkte für eine Öffnung des Wissenschaftsbetriebs und politisches Handeln. Als Beispiel des Empowerments nennen Corti und Fielding (2016: 7) Forschung mit First-Nation-Gemeinschaften in abgelegenen Teilen Kanadas. Bei dieser Arbeit werden Menschen der First Nation als Co-Forschende eingesetzt, sowohl direkt im Forschungsteam als auch an den entlegenen Orten, wo die Mitglieder der Community zum Feldteam werden. Das Forschungsteam versucht, die Interessen der Gemeinschaft nicht nur durch die Bereitstellung von Online-Ressourcen und Bildungsdiensten zu fördern, sondern auch durch die Vertretung seiner Interessen gegenüber regionalen und nationalen Regierungen.

Eine der Interviewten berichtete in diesem Zusammenhang, dass sie eine Archivierungsanfrage hinsichtlich einer Forschung zum Rohstoffabbau in Afrika und den dortigen Arbeitsbedingungen erhalten habe. Diese Daten könnten für Menschenrechtsorganisationen hohe Relevanz bekommen, wobei die Einrichtung den weiteren Umgang damit derzeit noch diskutieren muss. Als weiteres Beispiel nannte sie eine bereits durchgeführte Studie in Südamerika, wo eine gemeinsame Forschung mit einer Umweltorganisation durchgeführt wurde und diese nun ebenfalls Interesse an der Nutzung der gewonnenen Daten hat (vgl. TN2).

Ethische Fragen beim Management qualitativer sozialwissenschaftlicher Forschungsdaten sind oft komplex und fallspezifisch. Es ist stets ein Abwägungsprozess notwendig, der die Spezifik der Forschung im gesamten Forschungsverlauf miteinbezieht und sich nicht einmalig vorab klären lässt. Relevante ethische Fragestellungen entstehen immer wieder

während der Forschung, die „eine(r) forschungsethische(n) Reflexivität, die im Prozess sensibel auch auf unvorhergesehene Wendungen reagiert und das Forschungshandeln fortwährend einer kritischen Prüfung unterzieht“ (von Unger 2018: 92) nötig macht.

Forschende dürfen mit diesen Fragen nicht allein gelassen werden. Eine Möglichkeit könnte in der Etablierung von Data Access Committees bestehen, wie sie beispielsweise bei medizinischer Forschung bereits etabliert sind und die darüber entscheiden, ob eine Datennutzung zulässig ist.¹⁶ Es besteht die Notwendigkeit der Beratung, insbesondere, weil dahin gehende Entscheidungen nicht gänzlich an Rechtsexpert_innen/Ethikkommissionen abgegeben werden können und somit die Forscher_innen eigenverantwortlich entscheiden müssen. Zunächst müssen vor allem auf fachlicher Ebene Diskussionen über diese Fragen geführt und ein Bewusstsein dafür gefördert werden. Dieser Reflexionsprozess über Recht und Ethik sollte auch Teil der Lehre werden. Gleichzeitig sollte darüber nachgedacht werden, wie rechtliche und ethische Beratungsangebote effektiv in den Forschungsprozess integriert werden können, um Verunsicherung zu vermeiden (vgl. Huber 2019b: 39).

4.2) Fragen rechtlicher Natur

In Bezug auf die Archivierung, Veröffentlichung und Nachnutzung von Forschungsdaten sind zwei problematische Rechtsbereiche von zentraler Bedeutung: Zum einen ist zu klären, inwiefern das Urheberrecht relevant ist und falls dies der Fall ist, wer die entsprechenden Nutzungsrechte innehat. Zum anderen spielen datenschutzrechtliche Fragestellungen eine wichtige Rolle (vgl. Spindler/Hillegeist 2011: 64), wobei beide Felder bereits in der Planungsphase der Forschung Berücksichtigung finden müssen (vgl. Schönbrodt et al. 2017: 23). Mitunter sind jedoch auch, insbesondere bei internationalen Forschungskontexten, andere Rechtsvorschriften von Relevanz. So wurde in einem Interview erwähnt, dass ein Forschungsprojekt im Iran die Produktion von Fotos umfasste, wobei dann als Auflage des Forschungsdatenzentrums von den Datengebenden sichergestellt werden musste, dass das iranische Recht, welches das Fotografieren von Regierungs- und Militärgebäuden verbietet, eingehalten wird (vgl. TN2: 190-95).

¹⁶ Vgl. exemplarisch Kompetenznetz AHF: <https://www.kompetenznetz-ahf.de/wir/data-access-committee/>

4.2.1) Urheberrecht

Das Urheberrecht bezüglich digitaler Forschungsdaten ist komplex. So schreibt Kuschel: Die „Frage, ob überhaupt Rechte an ihnen bestehen, kann nicht pauschal beantwortet werden. Ein „Eigentum an Daten“ existiert (jedenfalls noch) nicht. Allerdings können Forschungsdaten unter Umständen urheberrechtlichen Schutz genießen“ (Kuschel 2018).¹⁷ Zentral ist hier die Frage nach der geistigen Schöpfungshöhe, die gegeben sein muss, damit ein urheberrechtlicher Schutz für die Daten gewährt wird (vgl. Spindler/Hillegeist 2011: 63). Urheberrechtlich geschützt ist immer nur die Form des Werkes, nicht der Inhalt. Dies meint, dass zwar die spezifische Darstellung als Ergebnis menschlicher Arbeit, beispielsweise in Form textlichen Beschreibungen (Transkripte, Notizen) oder bildlicher Dokumentation (Videoaufnahmen, Fotografien), geschützt werden könnte, jedoch nicht die Informationen selbst, die in dem Datum vorhanden sind. Gilt urheberrechtlicher Schutz für Forschungsdaten, müssen grundsätzlich Rechteinhaber_innen bei Fragen des Vervielfältigens, des Veränderns oder der Datenzugänglichkeit zustimmen (vgl. Kuschel 2018), wobei u. a. zu klären ist, wer nach Ende einer Forschung die Rechte innehat und über jene entscheiden darf (vgl. Gebel et al. 2015: 18).

Problematisch ist dies beispielsweise im Hinblick auf die verwendeten Einverständniserklärungen, aber auch bei älterem Material:

„Einerseits sichern in der qualitativen Sozialforschung eingesetzte Einwilligungserklärungen den Teilnehmenden an der Erhebung zumeist eine Nutzung der Forschungsdaten „nur im Rahmen dieses einen Forschungsvorhabens“ zu. Eine Weitergabe der Forschungsdaten ist damit ausgeschlossen, eine Archivierung zumindest erschwert. Andererseits lagern große Bestände historischer qualitativer Forschungsdaten für die Scientific Community unzugänglich bei Einzelforscher/innen und in Forschungseinrichtungen, da für diese Daten, begründet in einer älteren Forschungspraxis, keine schriftlichen Einwilligungserklärungen vorliegen.“ (Gebel et al. 2015: 4) Forschende können ihre Daten an Datenrepositorien mittels Datengeber_innenvertrag übertragen sowie mit Hilfe von Lizenzvergaben eindeutig deren Nachnutzung regeln. Die befragten Infrastruktureinrichtungen verwenden Nutzungsvereinbarungen, um so die Rechte und Pflichten von Nutzer_innen vertraglich festzulegen. Die Bereitstellung von Übergabe- und

¹⁷ Vgl. auch: „Eindeutige Klärungen rechtlicher Fragen sind derzeit jedoch nur begrenzt möglich, denn die Rechtslage ist mit Blick auf Forschungsdaten weitgehend „terra incognita“, für viele Felder gibt es noch kein Recht bzw. ist der Ausgestaltungsprozess im Fluss. Deshalb sind hier bestenfalls sukzessive Klärungen zu erwarten.“ (Imeri et al. 2019: 16) Eine abschließende Klärung des Sachverhalts ist an dieser Stelle deshalb nicht möglich.

Nutzungsvereinbarungen wurde von einer Interviewten als ein Bestandteil des Forschungsdatenmanagementprozesses beschrieben, in dem mit generischen Lösungen gearbeitet werden kann, da diese nur bei Bedarf ergänzt werden müssen (vgl. TN2: 617-32).

Zudem wurde darauf hingewiesen, dass insbesondere bei sogenannten digitalen Verhaltensdaten, die aus sozialen Medien gewonnen werden, eine intensive Beschäftigung mit den allgemeinen Geschäftsbedingungen der Plattformen notwendig ist, um zu klären, wie die Daten weiterverwendet werden. Die Entwicklung von dahin gehenden Standards stellt jedoch bislang noch ein Desiderat dar (vgl. TN7: 132-44). Eine Schwierigkeit besteht in der Abwägung „dafür geeignete Verfahren zu entwickeln [...], die aber auch nicht so speziell werden zu lassen, dass sie mit anderen Dingen nicht mehr kompatibel sind.“ (ebd.: 121-22)

Die Mitbestimmungsrechte der Forschungsbeteiligten werden in der Regel bereits vor der Archivierung durch eine Zustimmungserklärung zur Datenweitergabe vertragsrechtlich beschränkt (vgl. Huber 2019a: 17). Insbesondere bei Daten, die als zwischen Forscher_innen und Beforschten koproduziert angesehen werden, ist eine eindeutige besitzrechtliche Zuordnung oft nicht möglich und muss auch unter forschungsethischen Gesichtspunkten betrachtet werden.¹⁸ Dies umfasst ein „Spektrum von voller Eigentümerschaft über die eigenen Aussagen beforschter Personen(gruppen) (im Sinne von geistigem Eigentum oder Kulturerbe) über koproduzierte Daten (z. B. Daten, die im Rahmen von Interviews, Foto- oder Videoaufnahmen gewonnen wurden) bis hin zu Daten, die den Forschenden gehören (z. B. Notizen von Beobachtungen im öffentlichen Raum).“ (dgv 2018: 4) Hier wäre zu klären, inwiefern diese Kooperation die Miturheber_innenschaft an den Daten begründen könnte oder zumindest mit aktuellen Datenmanagementplänen transparent zu machen, welche dahin gehenden rechtlichen Absprachen mit den Beforschten getroffen wurden. Eine tatsächliche Kokuration von Daten ist derzeit vonseiten der Repositorien nicht möglich (vgl. Huber 2019a: 15, 17). Die Entwicklung weiterer maßgeschneiderter Rechtslösungen ist nötig, um mehr Transparenz und Rechtssicherheit sowohl auf Repositorienseite als auch für Forschende zu schaffen (vgl. Imeri et al. 2019: 32). In einem Interview wurde beschrieben, dass aus diesem Grund neue Arbeitsabläufe und Rechtslösungen angestrebt werden, um die

¹⁸ Vgl. hierzu: „Ownership: Ethnographic materials are coproduced by researchers and research participants and are embedded in particular social contexts. As such, they cannot be fully owned or controlled by researchers, research participants or third parties. The use of standard intellectual property licenses and protocols may not apply to all ethnographic materials.“ (EASA 2018: 1) Daraus folgt: „The first duty of anthropologists and ethnographers towards science is therefore to recognise this joint production and joint ownership of data. All forms and norms of managing data depend on it [...] Researchers should continue to treat data as collaborative for as long as they work with this material.“ (Pels et al. 2018b: 397)

Beforschten miteinzubeziehen und im Zuge dessen zu klären, wie Zugänge geschaffen werden können. Auch hier ist ein Abwägungsprozess nötig, „weil einerseits sollen die Daten ja geschützt sein und eben nicht so breit zugänglich sein. Und auf der anderen Seite hat man dann aber größere Gruppen, die Zugang haben sollen und sogar bis dahin, dass sie an den Daten was ändern dürfen“ (TN1: 469-71).

4.2.2) Daten-, Quellen- und Persönlichkeitsschutz

Neben den Fragen, wer originär die Rechte an den Daten innehat und wer diese auf welche Weise nutzen darf, ist ein weiteres Feld, das rechtlicher Klärung bedarf, die Frage nach den Rechten Dritter. Qualitative Forschung erzeugt oftmals Material, das datenschutzrechtlich ein erhöhtes Schutzbedürfnis aufweist, worunter die Verarbeitung „besonderer Kategorien personenbezogener Daten“ (EU 2016: 38) fällt. Die europäische Datenschutz-Grundverordnung definiert dies folgendermaßen:

„Die Verarbeitung personenbezogener Daten, aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen, sowie die Verarbeitung von genetischen Daten, biometrischen Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung einer natürlichen Person ist untersagt.“ (EU 2016: 38).

Die Verarbeitung ist nur möglich, wenn die „betroffene Person [...] in die Verarbeitung der genannten personenbezogenen Daten für einen oder mehrere festgelegte Zwecke ausdrücklich eingewilligt“ (ebd.) hat. Als dies umfasst Fragen des Daten-, Quellen- und Persönlichkeitsschutzes, die eine Einrichtung als „Dreh- und Angelpunkt“ (TN1: 67) der Datenvorbereitung beschreibt. Das Datenschutzrecht schützt „die informationelle Selbstbestimmung von Personen und gewährt dem von einer Datenverarbeitung Betroffenen bestimmte Rechte.“ (Kuschel 2018) Werden personenbezogene Daten archiviert muss immer überprüft werden, inwiefern Merkmale, die Rückschlüsse auf konkrete Personen ermöglichen tatsächlich notwendig sind (vgl. Spindler/Hillegeist 2011: 68).

Durch Anonymisierung oder Pseudonymisierung ist zu gewährleisten, dass personenbezogene Angaben so unkenntlich gemacht werden, dass die individuelle Zuordnung nicht oder nur mit überproportional großem Ressourceneinsatz möglich ist (vgl. Medjedović/Witzel 2010: 75). Es sind dabei verschiedene Ebenen der Anonymisierung zu

unterscheiden: So können direkte Merkmale der unmittelbaren Personenidentifizierung wie der Name oder Wohnort entfernt werden. Darüber hinaus gibt es indirekte Identifikationsmerkmale, bei denen durch sekundäre Informationen wie Alter oder ausgeübte Tätigkeiten, Personen individuell erkennbar werden können (vgl. Harbeck et al. 2018: 135). Diese Angaben können beispielsweise durch Vergrößerung z. B. bei Gebietsangaben, Zusammenfassung z. B. Schaffung von Altersklassen oder Löschung unkenntlich gemacht werden (vgl. Medjedović/Witzel 2010: 76).

Ersetzung von Merkmalen durch weitestgehend äquivalente Platzhalter und nicht durch beliebige Kombinationen aus Buchstaben und Zahlen mit dem Ziel das Auswertungspotenzial zu erhalten wird als Pseudonymisierung bezeichnet (vgl. Gebel et al. 2015: 14). Diese Informationen liegen in der Regel jedoch nicht in hoch strukturierter Form vor, wodurch die Anonymisierung erheblich anspruchsvoller wird (ebd.: 72). Bei sensiblen Themen müssen zudem andere Informationen, wie Örtlichkeiten, die im Kontext mögliche Hinweise zu einer Identifizierung geben könnten, unkenntlich gemacht werden (vgl. Schönbrodt et al. 2017: 23).

Die Offenlegung sensibler Details gilt auch für die Forschenden selbst, da deren Emotionen sowie Sinneswahrnehmungen in den Forschungsprozess einfließen und sie in ihren Forschungsdaten als Individuum, samt ihren Schwächen, persönlich identifizierbar werden (vgl. Breidenstein et al. 2013: 71; Huber 2019a: 4). Im Zuge ihrer Befragungen von Forschenden beschreibt Imeri, dass die Bereitschaft, Daten in ein Archiv zu übertragen, umso mehr schwindet, je stärker die Forschenden annehmen, dass ihre Person durch die Forschungsdaten identifizierbar wird (vgl. Imeri 2017: 174). Es gibt Befürchtungen, dass sich die eigene Praxis der Aufzeichnung und Befragung deutlich ändern muss. Es bestehen Bedenken gegen tendenzielle Selbstzensur, Zwang zur Anpassung und die Beschränkung von Offenheit und Freude am Experimentieren (vgl. Imeri 2019: 54 f.). Dies könnte zur Folge haben, dass für den Erkenntnisprozess wichtige Materialien wie Feldtagebücher, die die persönliche Gefühlswelt der Forschenden widerspiegeln können, nicht oder nur mit besonderen Regelungen zugänglich sein könnten (vgl. Harbeck et al. 2018: 130). Eine Einrichtung beschrieb, dass die Einführung der Datenschutzgrundverordnung zu einer erhöhten Sensibilisierung geführt habe. Es gebe jetzt ab und an Nachfragen von Beforschten, was mit den Daten passiert, jedoch keine permanenten Absagen. Dies gilt als Beleg, dass die „Vertrauensbasis, die bei solchen Erhebungen aufgebaut wird [...] sehr wichtig“ und in „ihrer Wichtigkeit zu schützen und zu bedenken“ (TN6: 361-63) ist.

In der alltäglichen Arbeit weisen die Einrichtungen den Primärforschenden die Aufgabe der Anonymisierung bzw. Pseudonymisierung ihrer Daten zu, wobei zukünftig auch auf unterstützende Tools zurückgegriffen werden soll (vgl. TN1; TN6). Das originäre Wissen der Forschenden ist zentral, da diese „über oder durch ihre Feldkenntnis überhaupt erst sagen können: Das hier ist ein sensibles Datum, was ein Außenstehender vielleicht gar nicht so entdeckt.“ (TN1: 102-103) Erschwerend kommt hinzu, dass auch fremdsprachiges Material produziert wird, bei dem die Archive nicht prüfen können, ob die Anonymisierung in ausreichendem Maße erfolgt ist. Dies kann nur über eine Erklärung der Forschenden abgesichert werden. Ein Archiv, das derzeit mit älteren Beständen arbeitet, anonymisiert die Daten bislang noch nicht und lässt sich von den Forschenden die Zusicherung geben, dass die Daten weiterverwendet werden können. Bei kleineren Gruppen ist eine umfassende Anonymisierung kaum möglich, da die Beforschten trotzdem identifizierbar bleiben. Manche der Beforschten forderten ebenfalls explizit als Person identifizierbar zu bleiben (vgl. TN3: 90). Mitzudenken ist zudem, dass was als schutzwürdig gilt, sich im Laufe der Zeit ändern kann und nicht alle Kontexte umfassend antizipiert werden können (vgl. Imeri 2018a: 75).

4.2.3) Anonymisierung und gestaffelte Zugänglichkeit durch die Einrichtungen

Die Aufgabe der Archive umfasst in diesem Zusammenhang einerseits die Beratung der Datengebenden, wie die Daten an die Einrichtung übergeben werden müssen, um sie adäquat archivieren zu können, sowie deren Prüfung und Korrektur. So beschrieb eine Interviewte: „Und wir gucken [...] auf Anonymisierung, auf: Sind die von der Qualität nachvollziehbar, verständlich? Bildet das, dass ab, was der Mensch uns gesagt hat, was er uns gibt? Ist es nachnutzbar, sinnvoll nutzbar?“ (TN1: 602-4) Neben allgemeinen Fragen stellen sich jedoch auch viele die konkrete Forschung betreffende Fragen, beispielsweise bezüglich rechtlich adäquater Einwilligungserklärungen (vgl. TN7: 48-56).

Die Prüfung betrachtet zum einen direkte Identifikationsmerkmale, aber ebenso das Zusammenspiel mehrerer Informationen (z. B. Tätigkeiten und geografische Angaben), die eine Identifizierung einer Person möglich machen würde. Ein Mittel sind verschiedene Varianten von Vergrößerungen: Werden beispielsweise detailliert Berufe beschrieben, greift ein Archiv auf die International Standard Classification of Occupations (ISCO) zur Vergrößerung der Angaben bei möglichst gleichzeitiger Erhaltung der Aussagekraft zurück.

Die flexible Anwendungsfallbezogenheit ist dabei wichtig zu berücksichtigen. Eine andere Person beschrieb:

„Wenn (ein) Berufspädagoge diese Daten haben will für seine Forschung. Dann muss der natürlich wissen, dass der Mensch Rettungssanitäter ist. Dem nutzt eine Vergrößerung nichts. An der Stelle wäre es dann so, dass wir gucken würden andere Vergrößerungen wie z. B. die Krankheit, wenn da noch steht er hatte einen Herzinfarkt. Dass man dann diesen Herzinfarkt vergrößert, sodass man die Kombination von Informationen damit so entschärft, sozusagen, dass daraus nicht eine Identifizierung möglich ist.“ (TN1: 143-48)

Aus der Spezifik des Materials ergibt sich in diesem Zusammenhang jedoch, dass eine umfassende Anonymisierung oftmals äußerst arbeitsaufwendig ist und eine (vollkommene) Anonymisierung, falls überhaupt möglich, die zur Sekundäranalyse notwendigen Kontextualisierungs- und Interpretationsmöglichkeiten mitunter massiv einschränken kann. Innerhalb qualitativer sozialwissenschaftlicher Daten sind viele Informationen enthalten, die eine Identifizierung von Personen möglich machen, selbst wenn sie formal mittels Entfernung direkter Identifikationsmerkmale anonymisiert werden. Anonymisierungsmaßnahmen, wie Schwärzen etc. können mit einem Informationsverlust einhergehen, deshalb wird oftmals auf mitunter ressourcenintensive Pseudonymisierung zurückgegriffen, falls sie vom Materialtyp überhaupt möglich ist (vgl. Hollstein/Strübing 2018b: 4). So merken Meier zu Verl und Meyer an:

„Auch wenn diese Elemente in den Publikationen punktuell z. B. durch Pseudonyme ersetzt werden können, würde eine „flächendeckende“ Pseudonymisierung a) zu einer Bedeutungsverschiebung führen, b) die Aussagekraft der Materialien verringern und c) zu einem Arbeitsaufwand führen, den die Ethnografin, die eigentlich immer nur punktuell pseudonymisiert, kaum neben ihrer Forschung bewältigen könnte.“ (2018: 87)

Daraus folgt, dass das „Datenschutzkonzept [...] sich daher nicht darauf beschränken (kann), den Anforderungen des Datenschutzes zu genügen, sondern es geht um eine Gratwanderung zwischen Datenveränderungen zum Schutz der untersuchten Personen und Organisationen und der Nutzbarkeit und Qualität der so veränderten Daten mit dem Ziel wissenschaftlich ertragreicher und valider Sekundärforschung. Und drittens muss der Aufwand für diese Maßnahmen berücksichtigt werden, damit das Gesamtziel nicht an nicht zu bewältigenden Bearbeitungsaufgaben der DatenhalterInnen scheitert.“ (Dunkel/Hanekop 2018: 41)

Aus diesem Grund lässt eine Einrichtung das Material unanonymisiert und verpflichtet die Nutzenden vertraglich bei Weiternutzung der Materialien für eine

ausreichende Anonymisierung zu sorgen und das Datenmaterial nach Abschluss der Arbeit zu vernichten. Der Grad der Anonymisierung ist nicht „sinnvoll entscheidbar“ (TN5: 111) und droht die Grundlage der Interpretation zu zerstören: „Man müsste dann [...] ganze Textpassagen rausnehmen und hätte am Ende ein Gerippe, das für eine Interpretation nicht mehr sinnvoll ist.“ (TN5: 108-10) In einem weiteren Interview wurden ebenfalls Fälle genannt, in denen die Beforschten über eine detaillierte und leicht recherchierbare Internetpräsenz verfügten, sodass eine Anonymisierung die Nutzbarkeit des Materials zu stark einschränken würde: „Da ist dann nur die Möglichkeit, dass man das Einverständnis einholt für die nicht perfekte, anonymisierte Nachnutzung und da müsste man dann halt auch mit einer stark reduzierten Anzahl an Interviews zum Beispiel dann rechnen.“ (TN2: 311-13)

In einem Interview wurde von mehreren gerade durchgeführten explorativen Nutzungsstudien berichtet, die Auskunft über mögliche Anonymisierungs- sowie Archivierungsmöglichkeiten bei Material mit größerer Nähe zu den Forschenden wie Feldnotizen und schwierig zu anonymisierenden Daten wie audio-visuellen Aufnahmen geben sollen (ebd.: 9-22). In einem Archiv sind einige Feldtagebücher zwar schon archiviert worden, aber bislang fehlt hier noch die breite Praxis (vgl. TN3: 56). Neben inhaltlichen Problemen erschweren die verwendeten Materialarten und technischen Entwicklungen die Anonymisierung noch zusätzlich. Während sich Transkripte mit geringerem Aufwand anonymisieren lassen, wurde die Anonymisierung von Ton- und Videoaufnahmen als sehr aufwendig beschrieben.¹⁹ Die individuelle Verpixelung von Personen sei mit hohen Kosten verbunden und einzelne Personen könnten trotzdem durch den Kontext identifizierbar bleiben: „können wir derzeit nicht leisten, müssten wir dann einfach sperren auch. Oder Intranetlösungen anbieten. Das ist eine Ressourcenfrage.“ (TN4 176-78) Eine andere Einrichtung problematisierte, dass durch technische Entwicklungen und die größere Datenverfügbarkeit intensiv darüber nachgedacht werden muss, inwiefern die eigene Anonymisierung der Daten ausreicht und ob „diese Anonymisierung auch so weit (reicht), dass niemand anderes durch das hinzuspielen von weiteren Informationen Dinge tun kann, die wir nicht wollen?“ (TN7: 410-12)

In einem Interview wurde zudem deutlich, dass der Datenschutz von Seiten der Infrastruktureinrichtung an seine Grenzen kommen kann, wenn „die Sensibilität oder das

¹⁹ Vgl. Smioski (2013: 18) für konkrete Maßnahmen der Anonymisierung von audio-visuellen Materialien und die damit einhergehenden problematischen Einschränkungen für die weitere Datenauswertung.

Bewusstsein der Forschungsteilnehmenden“ (TN2: 590-91) bezüglich Fotografien und Videoaufnahmen fehlt. Insbesondere durch die weltweite Verbreitung von sozialen Medien könnten die Forschenden selbst oftmals nicht kontrollieren, welche gemeinsam gemachten Fotos öffentlich sichtbar im Internet auftauchen. Dies könnte die aufwendigen Anonymisierungsbemühungen mitunter konterkarieren (ebd.: 606-11). Trotz dieser Probleme wurde mehrfach darauf hingewiesen, dass die vorgenommenen Sicherheitsmaßnahmen bislang funktionieren würden, sodass bisher noch keine Versuche festgestellt werden konnten, Daten zu deanonymisieren oder es zu missbräuchlicher Nutzung gekommen sei (vgl. TN6: 321-22; TN5: 125-126).

Um analytisch sinnvoll genutzt werden zu können, brauchen die Daten einen Kontext. Praktiken der Anonymisierung bewegen sich daher in einem Spannungsfeld, das zwischen ethischen und rechtlichen Einschränkungen einerseits und den methodischen sowie forschungspraktischen Anforderungen und dem Forschungsinteresse an eine sachgerechte Interpretierbarkeit sowie mitunter den Ansprüchen der Beteiligten andererseits auszubalancieren ist (vgl. Hollstein/Strübing 2018b: 5; Smioski 2013: 17). Je stärker anonymisiert wird, desto mehr leidet die Dichte des Materials. Eine befragte Person beschrieb dies als „Zielkonflikt [...] der Schutzbedarf der Befragten versus Analysepotenziale, das heißt, im Moment, wo wir Daten anonymer machen mit klassischen Verfahren zerstören wir Analysepotenziale. Das ist der klassische Widerspruch.“ (TN7: 380-82)

Der von den Fachgesellschaften geforderte Vorrang der Primärforschung bedeutet einen fallspezifischen Abwägungsprozess hinsichtlich des Umfangs der Weitergabe von Forschungsdaten, bei dem die Forschungsdurchführung und das Schutzbedürfnis aller in der Forschung Involvierten Vorrang besitzt. Opitz und Mauer stellen in ihrer Machbarkeitsstudie fest, dass diejenigen, die sich weigern, ihre Forschungsdaten zu veröffentlichen, den Schutz der Anonymität der Befragten und die datenschutzrechtliche Unbedenklichkeit als Haupthindernisse für die Weitergabe von Daten genannt haben. Unter den Forschenden, die der weiteren Verwendung ihrer Daten zustimmen, wurde dieser Themenbereich mit 27 % am häufigsten erwähnt (vgl. 2005: 16). Ein unbeschränkt offener Zugang zum Material ist durch den Quellenschutz kaum möglich, vielmehr muss ein gestaffelter Umgang mit Offenheit gefunden werden, um die Primärforschung nicht einzuschränken und Anreize für die Zugänglichmachung für bereits gesammeltes Material zu erzeugen. So wurde in einem Interview konstatiert, dass das eine „absolut große Barriere ist [...] die Angst, dass Daten

veröffentlicht werden, die nicht in die Öffentlichkeit gehören und dadurch der Feldzugang schwieriger wird.“ (TN1: 452-53)

In Abstimmung mit den Datengeber_innen können Zugriffsbeschränkungen festgelegt werden, die je nach Sensibilität der Daten, abgestufte Schutzstufen mit verschiedenen Zugriffsmöglichkeiten umfassen. Diese können von einschränkungsloser öffentlicher Nutzung via Onlinezugriff, freier Verwendung durch Wissenschaftler_innen in Form von Datenweitergabe, über Zugang nach Authentifizierungsmaßnahmen wie Nutzungsanträgen und Nutzungsbegrenzungen auf geschützte Räume vor Ort (Safe-Rooms) bis zu (temporären) Embargos reichen, wo keinerlei Verwendung möglich ist (vgl. Medjedović/Witzel 2010: 80 f.).

In Safe-Rooms lassen sich die Daten lediglich an lokalen Arbeitsplätzen nutzen, nur eingeschränkt bearbeiten und nicht kopieren (vgl. Imeri 2019: 55). Eine Sperrung der Daten für Dritte durch Embargos und eine langfristige Klärung von Zugriffsrechten stellen eine Möglichkeit dar, um datenschutzrechtlichen Anforderungen und dem Schutz von Personen und Forschungskontexten Genüge zu tun (vgl. EASA 2018; Klump 2010: 111). Bedenkenswert sind Versuche des National Anthropological Archives, das begonnen hat, offenere Methoden zur Beschränkung von Sammlungen zu erkunden. Diese Methoden würden sowohl die Interessen der Datengebenden als auch die Themen ihrer Forschung berücksichtigen; sie würden prüfen, ob der Zugang zu kulturell sensiblen Sammlungen selektiv den Herkunftsgemeinschaften gewährt werden kann und es ermöglichen, dass die Übergabevereinbarungen nach dem Ermessen der Archive (oder der einheimischen Gemeinschaften) in festgelegten Abständen, beispielsweise alle 10 oder 20 Jahre, überprüft (und vielleicht geändert) werden können (vgl. Leopold 2008: 8).

Ein vollkommen offener Zugang wie in Open-Data-Konzepten propagiert (vgl. Jacobs/Humphrey 2004: 28; The Royal Society 2012: 9) kann sich wahrscheinlich nur bei unkritischen Daten gewährleisten lassen.

Eberhard und Kraus weisen in diesem Zusammenhang auf einen weiteren notwendigen Abwägungsprozess hin: „selbst wenn der Zugang zu den Daten beschränkt ist, können bereits die Metadaten, sollen sie ihren eigentlichen Zweck erfüllen, kritische Informationen preisgeben. Werden andererseits die Metadaten so allgemein gehalten, dass sie nichts Personenbezogenes verraten, dann riskiert man einen Verlust an Datendichte bis hin zur Unbrauchbarkeit.“ (Eberhard/Kraus 2018: 46)

Die Einrichtungen sehen alle bestimmte Zugangsbeschränkungen für die Nutzung ihres personenbezogenen Materials vor. Bezüglich der Maximen von Open Data gilt für sie: „So offen wie möglich, aber so geschlossen wie nötig. Und dann muss man auch von unserer Seite aus rechtlich über einen Vertrag oder technisch einen Riegel vorschieben.“ (TN6: 294-95) Viele gewähren in der Hauptsache Wissenschaftler_innen Zugang und binden diese vertraglich durch Nutzungsvereinbarungen, die die Nichtweitergabe der Daten an Dritte und deren Löschung nach Projektende beinhalten (vgl. TN1; TN5; TN6).

Eine Einrichtung unterscheidet verschiedene Zugangsmöglichkeiten: Für Daten mit geringem Schutzbedarf ist nur eine Registrierung auf der Homepage nötig, um sie zu erhalten. Die darauffolgende Stufe erfordert zusätzlich die Angabe über den Verwendungszweck. Die Nutzung beschränkt sich auf wissenschaftliche Forschung und Lehre. Die höchste Schutzstufe sieht vor, dass Daten lediglich nach einer Einverständniserklärung der Datengebenden, die vom Datenarchiv eingeholt wird und nur nach Unterzeichnung eines Nutzungsvertrags weitergegeben darf (vgl. TN6: 432-49).

Für ein kleineres Archiv ist eine gestaffelte Zugänglichkeit derzeit noch ein mittelfristiges Ziel. Viele Daten, die bisher ins Repositorium eingespeist wurden, sind öffentlich und ohne Anmeldung zugänglich. Bei gesperrten Bereichen ist nur eine Vor-Ort bzw. Intranetnutzung möglich, was in der Praxis bislang jedoch noch nicht vorkam. Für sehr sensible Datensätze ist derzeit noch keine Infrastruktur vorhanden (vgl. TN3: 105-10). Ziel ist: „größtmögliche Zugänglichkeit mit größtmöglichem Schutz, also faire Daten.“ (TN4: 290-91)

Insbesondere die größeren Einrichtungen bieten Safe-Room-Nutzungen an, wo die Daten in einer virtuellen Arbeitsumgebung bereitgestellt werden. Dort dürfen keine eigenen elektronischen Medien mitgebracht werden und es existiert auch kein Internetzugang. Ergebnisse und Notizen der Analyse werden vor der Mitnahme vonseiten der Infrastruktureinrichtung nochmals kontrolliert. Nutzungsverträge regeln, welche Datennutzung zulässig ist und verbieten Deanonymisierungsversuche (vgl. TN2: 132-36). Embargos seien insbesondere für jüngere Forschende, die nur eine begrenzte Zahl an eigenen Forschungen durchgeführt haben, wichtig, um auf diese Weise ihr Forschungskapital im wettbewerbsgeprägten Wissenschaftssystem zu schützen. „Das ist ihr Schatz, auf dem sie sitzen und das natürlich noch länger für sich alleine nutzen wollen.“ (ebd.: 516) Forschende,

die bereits einen großen Teil ihrer Karriere absolviert haben, könnten mit der Datenfreigabe dagegen lockerer umgehen (ebd.).

Eine Einrichtung bietet Daten zudem in zwei verschiedenen Abstufungen an, wobei das erweiterte sensible Datenmaterial nur via Safe-Room genutzt werden kann (vgl. TN6: 373-80). Eine Sicherungskopie der virtuellen Arbeitsumgebung wird ergänzend noch zur erneuten Nutzung durch Forschende sowie als Kontrollmöglichkeit, falls Forschende Dinge veröffentlichen, die nicht publiziert werden dürfen, vorgehalten. Der hohe Grad an sozialer Kontrolle, der durch die Notwendigkeit des persönlichen Erscheinens im Forschungsdatenzentrum zusätzlich erhöht wird, gilt als Grundlage dafür, dass „man eigentlich Missbrauch weitestgehend ausschließen“ (ebd.: 411) kann. Als ergänzender Service wird der Zugang zu anderen Forschungsdatenzentren im Safe-Room angeboten, was es ermöglicht, eine Vielzahl an sensiblen Daten zu nutzen ohne, dass langwierige Reisen an verschiedene Standorte durchgeführt werden müssen (ebd.: 418-26).

Eine befragte Person wies daraufhin, dass Zugriffsbeschränkungen ein sinnvolles Schutzkonzept jenseits der eigentlichen Bearbeitung der Daten darstellen:

„Das heißt, man kann nicht nur an den Daten anonymisieren, sondern man muss, man kann Daten ja relativ schwach anonymisieren und muss dafür Schutz, den Schutz Drumherum erhöhen. Das wiederum ist ein Thema, wo wir Skalierungsprobleme kriegen, das heißt, in dem Moment, wo wir Daten nur vor Ort in geschützten Räumen zur Verfügung stellen, in geschützten und besonders überwachten Räumen zur Verfügung, haben wir natürlich einen Flaschenhals. Wir können nicht Hunderte von Leuten durch Schleusen betreuen.“ (TN7: 388-93)

Aus den Aussagen der Befragten wurde deutlich, dass es keine einheitlichen Standardlösungen gibt, sondern es immer an den Forschungskontext angepasste Formen der Zugänglichkeit geben muss, wo im besten Fall „Modullösung(en)“ (TN4: 186) mit einer Staffelung von Zugangsrechten andocken können.

4.2.4) Informierte Einwilligung

Ein zentrales Instrument der rechtlichen Absicherung bei personenbezogenen Daten stellt die informierte Einwilligung (informed consent) dar. Dieses Konzept macht die individuelle

und freiwillige Entscheidung zur Teilnahme an der Forschung von hinreichenden Informationen über die geplante Forschung abhängig (Medjedović/Witzel 2010: 70). Die Beforschten sollen dabei über Forschungsziele sowie die anvisierte Nutzung der Forschungsdaten informiert und ihre explizite Zustimmung dazu eingeholt werden. Sie sollen in die Lage versetzt werden die potenziellen Folgen der Datenproduktion und deren weitere Verwendung nachvollziehen zu können (vgl. Gebel et al. 2015: 9). Dazu wird vielfach eine von Forschungsteilnehmer_innen zu unterzeichnende schriftliche Einwilligung, dass Institutionen ihre personenbezogenen Daten archivieren und weitergeben dürfen, empfohlen (vgl. Spindler/Hillegeist 2011: 67).

In der Forschungspraxis ist die Umsetzung dieses Konzepts jedoch problematisch und geht mit offenen Fragen einher: Unklar ist, wie viel Information vorab gegeben werden muss. Das Einholen eines schriftlichen Einverständnisses ist gerade in sensiblen Bereichen kaum möglich und kann die Privatsphäre der Forschungsteilnehmer_innen verletzen oder sie gar in Gefahr bringen. Forscher_innen befürchten, dass die Notwendigkeit einer stark formalisierten Einwilligung die Vertrauensbeziehung zu den Beforschten und den Feldzugang stören könnte. Gerade bei Forschung zu sensiblen Bereichen könnte die umfangreichere Dokumentation und damit assoziierte Negativfolgen Forschung verunmöglichen (vgl. Hollstein/Strübing 2018b: 7).

Als ein Beispiel können die Forschungen von Dilger zu HIV/AIDS in Tansania fungieren, diese beschreibt er dahingehend, dass bereits die offene Benennung des Themas innerhalb von Interviewfragen mit der Stigmatisierung der Befragten als unmoralisch einhergehen könnte und so ein direktes Sprechen über HIV/AIDS oder gar Einverständniserklärungen, die dieses Thema umfassen, ethisch nicht vertretbar waren (vgl. Dilger 2017: 200 f.). Bezüglich der Institutionalisierung von Überprüfungsprozessen zur Einhaltung ethischer Standards im Forschungsprozess plädiert er für optionale sowie disziplinspezifische Orientierungs- und Bewertungshilfen für Forschende, die jedoch von Reflexivität und Kreativität anstelle von starrer Formalisierung und Vorschriften geleitet werden sollten (ebd.: 203 ff.).²⁰

Erschwerend kommt hinzu, dass die Anforderungen in unterschiedlichen Ländern variieren können und insbesondere bei Gruppen, speziell im ethnologischen Kontext, umstritten sein kann, wer legitimiert ist als Sprecher_in zu agieren (vgl. Eberhard/Kraus

²⁰ Vgl. hierzu auch Castillo 2018: 406 f.

2018: 48). Eine Interviewte wies darauf hin, dass diese Vorgaben in manchen Forschungskontexten deshalb problematisch wären, da sie eine ethnozentrische Sichtweise hätten, die von jeweils autonomen Individuen als Rechtssubjekten ausgeht und davon abweichende Konzeptionen zu wenig beachtet. Darüber hinaus gebe es Kontexte, in denen die Forschungsteilnehmenden explizit genannt werden wollen und ihr Recht auf Öffentlichkeit einfordern (vgl. TN3: 87-90).

Forschungsprozesse sind vielfach offen, es entstehen Daten auf Grundlage von Situationen und Dynamiken, die von Forscher_innen oft nicht antizipiert oder kontrolliert werden können (vgl. EASA 2018: 1). Eine im Voraus unterzeichnete formale Einverständniserklärung kann nicht alle Risiken und Eventualitäten des Wissenstransformationsprozesses abdecken (vgl. Pels et al. 2018b: 398). Zukünftige Sekundärnutzungen sind kaum umfassend antizipierbar. Insbesondere formale schriftliche Einwilligungserklärungen machen deshalb einen umfangreicheren Vorschuss an Vertrauen vonseiten der Beforschten nötig (vgl. von Unger 2018: 95). Es könnte eine abschreckende Wirkung auf Beforschte haben, wenn sie ihr Einverständnis nicht nur für eine konkrete Forschung geben sollen, sondern ebenfalls für Zukünftige, durch ihnen unbekannte Forschende (vgl. von Unger 2018: 97). Auch hier verschränken sich formalrechtliche und ethische Fragestellungen, für die es zeitlich und formal flexible Lösungsansätze braucht.

„Rechtliche Logik und forschungsethische Standards fallen (dabei) nicht unbedingt in eins. Denn während nach datenschutzrechtlichen Bestimmungen nur ein Einverständnis für die Speicherung, Verwendung und Weitergabe personenbezogener Daten erforderlich ist, kann dies aus forschungsethischer Perspektive auch für Daten ohne Personenbezug – etwa mit Blick auf ganze Gruppen oder auch für anonymisierte Daten gelten.“ (dgv 2018: 3)

Der Umgang damit lässt sich nicht auf das Format von im Voraus einzuholenden schriftlichen Zustimmungserklärungen beschränken (vgl. ebd.). Das Geben des Einverständnisses kann oftmals nicht standardisiert vereinbart werden und muss deshalb als „ongoing process“ (EASA: 2018: 1), der flexibel an das Forschungsfeld angepasst wird, konzipiert werden. Im Falle von audio-visuellen Daten schlagen Pels et al. (2018b: 399) vor: „a common practice is to follow up initial consent by people studied with a reaffirmation by their viewing of the finished product, and by negotiations about ownership of and access to the recordings, conditions of publication and copyrights of the eventually published film.“

In diesem Zusammenhang ist es hilfreich, dass die Datenschutz-Grundverordnung (DSGVO) zudem einen sogenannten broad consent ermöglicht, der eine

Einwilligungsmöglichkeit darstellt ohne, dass die Forschungszwecke final konkretisiert sind. Dadurch ist es rechtlich gestattet, Daten auch dann nutzen zu dürfen, wenn sie vorab für andere Zwecke erstellt wurden, wobei ethische Einschränkungen immer auch mitreflektiert werden müssen (vgl. Kuschel 2018). Das Geben von Einverständnis müsste in diesen Zusammenhang eher als ein andauernder und personengebundener Aushandlungsprozess, basierend auf in der Ausbildung und Forschung angeeigneten Ethikstandards begriffen werden, der sich im Verlauf der Forschung wandeln kann (vgl. dgV 2018: 3; Imeri et al. 2019: 15). Eine Interviewte wies auf die Notwendigkeit hin, dass sich die betroffenen Wissenschaftsdisziplinen verstärkt für datenschutzrechtliche Sonderregelungen einsetzen sollten, die Einwilligungsprozesse vereinfachen würden, sodass eine mündliche Einwilligung oder bezeugt durch Anwesenheit Dritter ebenfalls ausreichen würde (vgl. TN2: 555-60). Mit Hilfe einer Anwendungsstudie versucht die Einrichtung derzeit zu eruieren, inwiefern ein Umgang darin bestehen könnte, die informierte Einwilligung erst nach Produktion der Forschungsdaten, mittels eines Gesprächs, einzuholen, wobei es vorteilhaft wäre, dass zu diesem Zeitpunkt vonseiten der Beforschten ein höheres Maß an Vertrauen vorhanden sei (ebd.: 81-91).

In mehreren Interviews wurde ein pragmatischer Umgang mit Einverständniserklärungen deutlich. Ein kleineres Archiv beschreibt seinen Umgang mit dem Thema als bislang „eher unbeschwert[er]“ (TN3: 80). Dadurch, dass das Material schon älter ist und durch die Nähe zu Forschenden glaubt man diesen, wenn sie sagen, dass es datenschutzrechtlich in Ordnung ist. Bei geplanter forschungsbegleitender Datenarchivierung wird dies jedoch nicht mehr in diesem Maße möglich sein. Es wird darüber nachgedacht, eine Archivlogik zu implementieren, wonach die Schutzfrist für Material 100 Jahre nach der Geburt der betroffenen Person endet, wobei sich dies, durch den hohen Aufwand, nur in Einzelfällen realisieren lassen wird (ebd.: 80-86, 91).

Eine andere Einrichtung geht davon aus, dass die Datengebenden die Daten selbst produziert haben, und formuliert dann einen Übernahmevertrag, mit dem das Archiv von der Verantwortung freigestellt wird, „dass diese Nutzung im Archiv vielleicht gar nicht gewünscht war. In der Praxis kommt es eigentlich [...] überhaupt nie zu Problemen.“ (TN5: 125-26) Zudem sei es so, dass selbst wenn Einverständniserklärungen vorhanden sind, diese mehr beinhalten als diese tatsächlich abdecken können. Oftmals würden Dritte erwähnt werden, „deren Rechte gar nicht eingeholt werden. [...] Das halte ich für ein nicht unerhebliches Problem.“ (ebd.: 141-43) Gleichzeitig tritt im Zuge technischer Neuerungen

das Problem auf, dass sie durch viele ältere Einverständniserklärungen überhaupt nicht abgedeckt sind und z. B. die Onlinenutzung nicht antizipieren konnten. Das Archiv interpretiert die Erklärungen deshalb dahingehend, dass die Archivierungserlaubnis auch eine authentifizierte Onlinebereitstellung umfasst (ebd.: 128-36).

In einem anderen Gespräch wurden die Einverständniserklärungen als „vorsichtig gehalten“ (TN6: 256) beschrieben, wo weniger auf „bestimmte technische Möglichkeiten abgehoben“ (ebd.: 260) wird, da diese sich ändern können. „Es müssen halt die Bedingungen weiterhin gegeben sein, dass zum Beispiel ein Rückschluss auf die Person nicht möglich ist.“ (ebd.: 266-67)

4.2.5) Zwischenfazit ethische und rechtliche Fragen

Unsicherheiten im Hinblick auf ethische Probleme und eine unklare Rechtslage erschweren Forscher_innen den weiteren Umgang mit den Forschungsdaten und können dazu führen, dass diese auf eine Publikation ihrer Daten verzichten (vgl. Kaden 2018: 6; Queckbörner 2019: 65). Einige Fragen lassen sich auf formalrechtlicher Ebene schwierig abschließend lösen, sondern müssen immer im Verbund mit forschungsethischen Überlegungen und den Bedürfnissen der Beforschten betrachtet werden. Deshalb erfordern die beschriebenen Probleme die Entwicklung adäquater Standards sowie flexible und projektspezifische Lösungsansätze, um so sicherzustellen, dass Unsicherheit verringert wird und Daten nicht verloren gehen. Eine „Überregulierung und Bürokratisierung der Prozesse“ (dgv 2018: 2) ist, durch die vielfach dynamische Entwicklung der Feldzugänge sowie des Forschungsfokus, problembehaftet und soll mit Blick auf die Forschungspraxis vermieden werden. Vorab erstellte Datenmanagementpläne können deshalb in Hinblick darauf, welche Daten, in welchem Umfang und mit welchen Personen erzeugt werden, nur eine vage Richtschnur liefern und müssen auf unvorhersehbare und unkontrollierbare Ereignisse im Forschungsprozess reagieren (vgl. Lederman 2018: 408).

Generische Regelungen sind in den meisten Fällen nicht zielführend und eine projektbegleitende Beratung, beispielsweise in Hinblick auf Einwilligungserklärungen, Anonymisierungsmöglichkeiten sowie die Ausgestaltung von Nutzungsrechten von Infrastruktureinrichtungsseite, sinnvoll. Huber merkt an, dass es aufseiten der Universitäten häufig an spezifischen Lösungen mangelt. Die vorhandenen Angebote für Rechts- und

Ethikfragen beschränken sich oft auf den Datenerhebungs-, Nutzungs- sowie Publikationsprozess, eine Leerstelle ist vielfach bei der Frage der Selektion und Aufbereitung von Forschungsdaten für Sekundärnutzungen zu konstatieren (vgl. Huber 2019b: 36 f.). Eine Einrichtung äußerte, dass es sinnvoll wäre, wenn Ethikkommissionen versuchen könnten, Orientierungshilfen speziell für ethnografische Methoden zu schaffen (vgl. TN2: 560-62). In einem Interview wurde hervorgehoben, dass es von großer Wichtigkeit ist, Fragen in diesem Zusammenhang frühestmöglich zu klären und potenzielle Datengebende dafür zu sensibilisieren, da nur dann die Einrichtung, beispielsweise in Hinblick auf die Ausformulierung von Einverständniserklärungen oder Urheberrechtsfragen, mit ihrem reichhaltigen Erfahrungswissen beratend tätig werden kann (vgl. TN6: 227-44).

4.3) Praktisch-technische Probleme im Umgang mit qualitativen sozialwissenschaftlichen Forschungsdaten

Neben rechtlich-ethischen Fragen existieren auch praktisch-technische Herausforderungen, die sich aus der Spezifik der qualitativen sozialwissenschaftlichen Forschungsdaten ergeben. Die Heterogenität der Forschungsmethoden führt dazu, dass die Daten wenig standardisiert und multimodal sind. Sie können beispielsweise Texte, Fotos, Audio-Visuelles, Karten sein, wobei Daten in Textform (z. B. Transkriptionen oder Protokolle) am weitesten verbreitet sind (vgl. Imeri et al. 2019: 12).

Problematisch ist, dass es für die Bearbeitung keine Standardsoftware gibt und dass Datensets mitunter in proprietären technischen Formaten gespeichert werden, die bisweilen nicht kompatibel mit freien Softwarelösungen sind. Gewerbliche Softwarelösungen erzeugen oft proprietäre Sonderformate, die ausschließlich mit einer gültigen Lizenz wieder geöffnet und bearbeitet werden können. Beispielhaft kann hier die häufig verwendete Analysesoftware MAXQDA erwähnt werden, in der bearbeitete Datensätze im proprietären .mx-Format gesichert werden. Wenn diese weiterverwendet werden sollen, ist man zwangsläufig auf die Nutzung dieses, mit hohen Lizenzgebühren einhergehenden, Programms angewiesen (vgl. Sterzer/Kretzer 2019: 113). Geht es in der Hauptsache um Material in digitaler Textform, könnte eine Möglichkeit darin bestehen dieses in Standardformate wie XML zu überführen, was gegenwärtig als adäquat für vielfältige Nachnutzungen und eine dauerhafte Archivierung erscheint (vgl. Queckbörner 2019: 41).

Archivierungsworkflows müssen primär an die Forschungsabläufe angepasst werden, Sterzer und Kretzer empfehlen deshalb, dass „alle gängigen Formate, die im Forschungsalltag durch diverse Software und Hardware entstehen können, Berücksichtigung finden.“ (2019: 113) Einschränkungen im Hinblick auf Datenformate könnten zur Folge haben, dass die Methoden der Datenerhebung mitunter zu stark eingegrenzt werden könnten, was die Übergabebereitschaft an Forschungsdatenrepositorien verringern kann (Imeri et al. 2019: 25). Technisch flexible Infrastrukturen, die Datenbeziehungen detailliert abbilden können, sind ebenso notwendig wie ein hohes Maß an fachspezifischer Spezialisierung (vgl. Harbeck et al. 2018: 137). In den Interviews wurde eine prinzipielle Offenheit für viele Daten(-formate) in der Arbeitspraxis deutlich, die wenige Beschränkungen beinhaltet (vgl. TN1; TN3; TN5; TN6).

Der methodische Pluralismus bei qualitativen sozialwissenschaftlichen Daten bedeutet zusätzlichen Arbeitsaufwand und wirft Fragen hinsichtlich der Darstellung durch Metadaten, der Etablierung von Qualitätsstandards und der Prozessierung der Daten auf (vgl. TN1: 629-31). Eine Einrichtung, die sich mit Daten aus sozialen Medien beschäftigt, problematisierte die damit einhergehenden großen Datenmengen. Diese stellen neue Anforderungen an die technische Infrastruktur. Datensätze im Terabytebereich: „können wir nicht mehr einfach so von A nach B schubsen oder irgendjemandem zum Download anbieten.“ (TN7: 161-162) Selbst, wenn die Datenweitergabe funktioniert, mangelt es den Nutzenden wiederum an Rechenkapazität für eine weitere Verarbeitung.

Gerade bei älteren Datenbeständen erfordert der Medienbruch zwischen analogen und digitalen Beständen ebenfalls Lösungen. Analoge Forschungsdaten sind insbesondere im Kontext der Archivierung von Nach- und Vorlässen von Wissenschaftler_innen von Relevanz (vgl. Imeri et al. 2019: 42). So beschreibt Smioski, dass die meisten Datensätze von Forscher_innen, die ihren Ruhestand planen, stammen. Der drohende Datenverlust ist in diesem Fall ein wesentlicher Faktor für die Archivierung (vgl. Smioski 2011: 224). Hinzu kommt, dass beispielsweise in Feldforschungssituationen vielfach nach wie vor analoges Datenmaterial, zum Beispiel Archivmaterial, genutzt wird bzw. das Feld die Aufzeichnung mittels analoger Methoden erfordert (vgl. Imeri 2019: 53).

In einem Interview wurde von der Archivierung von 120 Ordnern mit analogem Datenmaterial samt handschriftlichen Notizen berichtet, die gescannt wurden. Hier wurde der große Aufwand problematisiert, da zu einer wirklichen Nutzung neue Transkriptionen nötig wären, die von den Sekundärforschenden zu leisten wären, da die technischen

Lösungen für eine automatisierte Texterkennung derzeit für dieses Material noch nicht ausreichen würden (vgl. TN1: 202-13).

Eine andere Einrichtung, die viel mit Vor- und Nachlässen von Forschenden arbeitet, hat optimierte Digitalisierungsworkflows und Dateienbenennungsregeln für ihre Bestände entwickelt. Die Entwicklung der Prozesse ist sehr aufwendig und sieht im Anschluss ein Eingabeformular für das Repositorium vor, in dem die Daten dann noch mit Metadaten versehen werden, bevor sie schließlich validiert und zugänglich gemacht werden. Parallel wird, falls nötig, mit den Forschenden kommuniziert, um eventuelle Unstimmigkeiten z. B. uneinheitliche Bezeichnungen in den Metadaten zu klären (vgl. TN4: 61-77). Aufgrund mangelnder Personalressourcen soll zukünftig Forschenden die Möglichkeit der Selbsteingabe durch eigene Accounts ermöglicht werden, wobei deren Betreuung trotzdem mit viel Aufwand verbunden ist. Zugleich geht dies für die Forschenden ebenfalls mit sehr viel Arbeit einher, die eine intensive Zusammenarbeit mit dem Archiv und eine eingehende Auseinandersetzung mit den eigenen Daten notwendig macht (vgl. TN4: 139-45).

Die bislang als open-source verfügbaren Archiv- und Toollösungen sind häufig zu generisch und besitzen zu wenig flexible Funktionen. Sie können zur Unterstützung spezifischer qualitativer Workflows oftmals nicht breit angewandt werden, vielfach benötigen sie mitunter arbeitsaufwendige Anpassungen (vgl. dgv 2018: 2; Imeri et al. 2019: 23). So kommentieren Harbeck et al.: „In der Regel zeigen sich die Betreiber der Datenarchive skeptisch hinsichtlich der funktionalen Kapazitäten vorhandener open-source Archivlösungen wie z. B. DSpace. Sie sehen Probleme, etwa bei der Implementierung eines gestaffelten Rechtemanagements, bei der Herausforderung, einzelne Dateien innerhalb von Projekten mit unterschiedlichen Beschreibungen zu versehen, oder berichten über unzureichende Kuratierungsfunktionen für heterogene Datenformate.“ (2018: 134)

Eine praktikable generische Möglichkeit, um den Umgang mit qualitativen Daten zu organisieren, sind Datenmanagementpläne. In einem Interview wurde in diesem Zusammenhang darauf hingewiesen, dass die „Akzeptanz ist ja relativ gering gewesen, die zu nutzen“ (TN2: 618), außer, wenn dies von der Forschungsförderung her gefordert wird. Um Verunsicherung hinsichtlich Strukturierung und Gestaltung zu vermeiden und Anreize zu schaffen wurde vorgeschlagen, dass besonders gelungene Managementpläne zukünftig veröffentlicht werden (ebd.: 619-24) und diese somit nicht als aufgezwungene Maßnahme angesehen werden.

In einem anderen Interview wurden Einschränkungen hinsichtlich der Realisierung vorab erstellter Datenmanagementpläne, die das Archivierungs- und Nachnutzungspotenzial präzise bestimmen, aufgrund der Offenheit von Forschungsprozessen deutlich (ebd.: 616-24). Dies wurde anhand einer Erprobungsstudie beschrieben: In der beforschten Region kam es zu politischen Unruhen, was zur Einschränkung der Bewegungsfreiheit und verstärkter Archivforschung führte. In Abweichung zu den Plänen des Forschungsantrags konnten weniger Interviews geführt und Personen nur so fotografiert werden, dass die Anonymität gewahrt wurde. War die Forscherin anfänglich der Weitergabe der Feldnotizen gegenüber noch sehr offen eingestellt, wurde im Laufe der Forschung klar, dass diese zu persönlich sind. Dies bedeutete, dass diese nicht direkt archiviert werden können, sondern daraus zuerst Memos produziert werden müssen, die als erläuterndes Kontextualisierungsmaterial den Forschungsdaten, zwecks Beschreibung, beigelegt werden können. Es stellte sich heraus, dass weniger archiviert werden konnte und Teile der Daten aus Selbstschutz vorerst mit einem Embargo belegt werden mussten oder nur mit ihrer Einwilligung anderen Forschenden zugänglich sind (ebd.: 40-75).

4.3.1) Erschließung der Forschungsdaten

Um archivierte Daten nutzen zu können, müssen diese zuerst erschlossen werden. Hierbei sind die detaillierte Beschreibung durch dokumentierende Metadaten auf Grundlage von kontrollierten Vokabularen, darüber hinausgehende Kontextualisierung sowie der Rückgriff auf persistente Identifikatoren, relevant.

1.) Dies ermöglicht, dass Forschungsdaten nachhaltig erschlossen und Jahre nach der ursprünglichen Entstehung nachgenutzt werden und potenziell als eine für sich stehende Publikation angesehen werden können (vgl. Jensen et al. 2011: 83; Schönbrodt et al. 2017: 27). Einerseits erleichtern stark auf Forschungsgebiete zugeschnittene Metadatenschemata es, zentrale Informationen aufzunehmen und auszutauschen - eine Übertragbarkeit auf andere Gebiete ist mitunter jedoch schwierig zu realisieren. Um eine breite Nachnutzung zu ermöglichen, muss zwischen generischen standardisierten Beschreibungen und solchen, die der Spezifik und den Kontext der Daten erhalten, abgewogen werden (vgl. Dunkel/Hanekop 2018: 40). Durch die nicht-standardisierten Methoden der qualitativen Datenerhebung sind flexible Metadatenschemata, je nach Materialart, nötig. Dies führt vielfach dazu, dass die

Repositorien mehrere Metadatenschemata zur Beschreibung ihrer Daten nutzen und spezifische Standards entwickeln (vgl. Harbeck et al. 2018: 130; Imeri et al. 2019: 23 f.).

Für Studien z. B. im UK Data Archive wird dabei ein systematischer Metadaten-Datensatz erstellt, der einen Überblick über die Studie, die Größe und den Inhalt der Datenbestände sowie die Zugangsbedingungen bietet (vgl. Corti 2018: 22). Im Bereich der sozialwissenschaftlichen Daten gibt es das ursprünglich für quantitative Daten entwickelte DDI-Format der Data Documentation Initiative (DDI) als Metadatenschema. Es enthält einen umfangreichen Satz beschreibender Elemente, darunter den Titel, Datengebende, Zusammenfassungen, Erhebungsmethoden und Forschungszeitraum, Urheberrechte etc. (vgl. Jensen et al. 2011: 87). Dieses Schema ermöglicht den Datenaustausch mit weiteren Archiven mittels übergreifender Datenverzeichnisse (vgl. Smioski 2011: 229). Imeri et al. weisen jedoch darauf hin, dass innerhalb ihrer Erhebungen häufiger erwähnt wurde, dass dieses Metadatenschema weitgehend ungeeignet für das Erfassen von audio-visuellen Materialien ist (Imeri et al. 2019: 25).

Bei einer Einrichtung sind die verwendeten Metadatenschemata insofern generisch, als dass sie an internationale Standards angepasst sind und es einige Pflichtfelder gibt. Durch die Nutzung von Freitextfeldern sind die Schemata jedoch zugleich spezifisch anpassbar (vgl. TN2: 624-627). Eine andere Interviewte beschrieb, dass sie ein eigenes umfangreiches Metadatenschema für ihre Daten entwickelt habe, das durch viele Freifelder auch auf verschiedene Fachspezifika anpassbar sei (vgl. TN1: 279-83). Forschungsrelevante Angaben werden durch die Forschenden selbst mit Hilfe eines „dialogische(n) Formulars“ (TN1: 286) eingetragen, zusätzlich als „Minimalkontextualisierung gibt es einen sogenannten Studienreport [...] (der) Themenbereiche (wie) Sampling, [zur] Methode, ein bestimmtes Abstract“ (ebd.: 287-89) und den Umfang des gesamten Materialkorpus umfasst, wobei der Report via Links wieder mit den Metadaten verbunden ist. Metadaten, die einzelne Fälle (z. B. ein konkretes Interview) innerhalb einer Studie beschreiben, sind, da potenziell sensibel, nicht öffentlich einsehbar. Die Metadaten sind interoperabel, was den Vorteil hat, dass sie von „jedem geharvestet werden“ (ebd.: 297) können und so eine internationale Repräsentation auf Datenplattformen möglich ist. Eine andere Interviewte beschrieb, dass gemeinsam mit den Forschenden herausgearbeitet wird, welche Metadaten für den Datensatz relevant sein könnten (vgl. TN3: 73-74). Eine weitere Einrichtung fordert von den Datengebenden die Abgabe von Metadaten durch ein Onlineformular ein und ergänzt diese Angaben durch eine eigene Materialerschließung (vgl. TN6: 148-60).

Als problematisch wurde beschrieben, dass anfänglich die Tendenz bestand, die Metadaten zu überfrachten (vgl. TN3: 35) sowie, dass Metadaten keine Autor_innenschaft besäßen und „immer sichtbar“ (ebd.: 37) seien. Da in diesem Archiv mit analogen und digitalen Beständen gearbeitet wird, wurde ein zweistufiges Metadaten-System entwickelt, womit sowohl das Objekt selbst, Fotos davon und zusätzliches Material wie Notizen als sogenannte Containerobjekte erfasst werden. Ziel ist es, anstelle von kleinen Dateneinheiten so eine übergeordnete und verknüpfende Kollektionseinheit zu schaffen, in der Materialzusammenhänge bewahrt werden (ebd.: TN3). An anderer Stelle wurde problematisiert, dass die Metadaten einerseits das Material mittels Abstraktion und Zusammenfassung auf verschiedenen Plattformen sichtbar machen sollen, andererseits jedoch „auch die Partikularität jeder einzelnen Forschung zu bewahren (ist). Da stehen wir noch davor, den Mittelweg zu finden.“ (TN2: 571-72)

Die weiteren Erschließungsansätze unterschieden sich je nach Einrichtung: Nachdem Transkriptionen lange als die einzige sinnvolle Erschließungsmethode für Interviews angesehen wurden, setzt ein Archiv nun auf automatisierte Spracherkennung. Dies funktioniert für qualitativ hochwertige Tonaufnahmen bereits sehr gut, hat bei Altmaterial minderer technischer Qualität derzeit noch seine Schwächen (vgl. TN5: 357-63). Das Archiv verzeichnet jedes Datum einzeln und gibt Auskunft über den Entstehungszusammenhang, die Form sowie ergänzende Informationen zur Interviewtenbiografie. Bei vielen Interviews bleibt die inhaltliche Erschließung jedoch, bedingt durch Ressourcenmangel, unvollständig. Falls Nutzende beim Durchhören der Interviews relevante Informationen finden, werden sie gebeten, diese in die entsprechenden Felder nachzutragen. Zusätzliche Transkriptionen kann die Einrichtung nicht erstellen, falls vorhanden, dienen diese Volltexte dann als zentrale Recherchegrundlage (vgl. TN5: 194-207). Ein großes Problem wird insbesondere darin gesehen, herauszufinden, welche Interviews für die Forschungsfrage von Relevanz sein könnten. Zwar werden vonseiten der Einrichtung aus, gemeinsam mit Forschenden, Suchbegriffe entwickelt, die Ergebnisse liefern können. Es bleibt jedoch unklar, ob besonders wertvolle Inhalte überhaupt gefunden werden. Hier erhofft sich die Institution zukünftig technologische Weiterentwicklungen durch Spracherkennungssoftware, sodass der Inhalt des Audiomaterials umfangreicher erschlossen werden kann und eine transkriptionslose direkte Suche im akustischen Signal möglich wird (vgl. TN5: 328-37).

2.) Ein wichtiger Bestandteil einer standardisierten Erschließung sind kontrollierte Vokabulare, die beispielsweise Hinweise zur Forschungsmethodik und dem -thema geben (vgl. Jensen et al. 2011: 88). Eine interviewte Einrichtung greift dabei auf die reduzierte deutschsprachige Version des European Language Social Science Thesaurus (ELSST) zurück, was zur groben Schlagwortvergabe bezüglich der Studienthemen genutzt wird (vgl. TN6: 186). Der eigens entwickelte Thesaurus wird aufgrund seiner Feingliedrigkeit als nicht gut geeignet für eine Inhaltserschließung betrachtet, „weil das immer sehr aufwendig ist die Thesaurusmanagementsysteme nachzuhalten.“ (ebd.: 173-74)

Mehrfach wurde problematisiert, dass ein kontrolliertes Vokabular „noch eine ziemliche Dunkelkammer bei qualitativen Daten“ (TN1: 283-84) sei. Die vorhandenen Fachvokabulare – sofern sie überhaupt existieren – und Namendatenschemata sind oft unbrauchbar. Die Etablierung von kontrolliertem Vokabular ist daher notwendig (vgl. TN3: 114). Eine auf ethnologische Daten spezialisierte Einrichtung ergänzt dieses beispielsweise durch die Erstellung von eigenen Vokabellisten „mit den wichtigsten ethnischen und geografischen Namen von Gegebenheiten“ (TN4: 81; vgl. auch Harbeck et al. 2018: 130). Die GND wurde in diesem Kontext als „eine Katastrophe“ (TN4: 95-96) beschrieben, weil „sie einfach hoffnungslos veraltet ist und teilweise noch wertende Bezeichnung drin hat [...] komplett unvollständig, riesige Lücken, Fehler.“ (ebd.: 96-98) Unklar sei bei ethnischen Bezeichnungen, welche gemeinsamen Standards für Transliterationen und Selbstbezeichnungen genutzt werden könnten (vgl. TN4: 98-102). Um weiter arbeiten zu können, fehlten Standards und „Thesauri ganz dringend, die brauchen wir und das kann keine Institution alleine stemmen.“ (ebd.: 316-20) Die GND müsste entsprechend aktueller Entwicklungen und Fachdebatten überarbeitet und ergänzt werden. Sinnvoll ist es darüber hinaus, dass in Repositorien bestenfalls später noch weitere Vokabulare sowie Metadatenschemata implementiert werden können (vgl. Imeri et al. 2019: 25). Forschungsdaten existieren mitunter in unterschiedlichen Sprachen, was auf Ebene der Bereitstellung und Nachnutzung eine Rolle spielt (vgl. TN2).

3.) Daneben erfordert die dauerhafte eindeutige Referenzierbarkeit der Daten und dadurch die Möglichkeit diese sichtbar, aufrufbar und zitierbar zu machen, den Einsatz von persistenten Identifikatoren. Dadurch bestehen Verknüpfungen dauerhaft und Datengebende können eindeutig identifiziert sowie eventuell kontaktiert werden (z. B. via ORCID) (vgl.

Corti/Fielding 2016: 3 f.; Queckbörner 2019: 58). Die interviewten Einrichtungen greifen dabei auf Digital Object Identifier (DOI) zurück (vgl. TN4: 248; TN6: 458-60).

4.) Vielfach geht es in der qualitativen sozialwissenschaftlichen Forschung darum, Handlungen oder Aussagen in ihrem Sinnzusammenhang zu verstehen. Forschende, die eine Sekundäranalyse durchführen, waren nicht Teil der ursprünglichen Situation und ihnen fehlt deshalb das unmittelbare Kontextwissen (vgl. Hollstein/Strübing 2018b: 4 f.). Qualitative Daten brauchen, um nachgenutzt werden zu können, deshalb oftmals ein hohes Maß an Datendokumentation in Form von Kontextualisierung, das nicht allein über Metadaten gesichert werden kann (vgl. Harbeck et al. 2018: 130). Damit soll der institutionelle Hintergrund, der theoretisch-methodische Studienansatz, der Datenerhebungs-, Aufbereitungs- und Analyseprozess, forschungsstrategische Entscheidungen, Überlegungen der Primärforschenden sowie Nutzungsbedingungen nachvollzogen werden können (vgl. Smioski 2011: 228 f.).

Bambey et al. unterscheiden hinsichtlich der Kontextualisierung drei Bereiche: Makro-, Mikro- und Objektebene. Die Makroebene gibt Auskunft über Studienhintergründe und Erstellungskontext (Projektbeteiligte, Finanzierung etc.), die in der Forschung verwendeten Methoden und Forschungsprozesse (z. B. Feldzugänge, Zeit und Ort der Forschung) sowie Rechtsaspekte (z. B. Nutzungsregelungen, Einverständniserklärungen). Die Mikroebene umfasst das Setting der Forschung: Sie kann zusätzliche Informationen über die soziale Einbettung beispielsweise einer Interviewsituation oder weitere Hintergrundinformationen bezüglich der Beforschten umfassen (z. B. Geschlecht, Alter sowie Bildung oder sozio-ökonomischer Hintergrund). Die Objektebene umfasst schließlich die Art und Weise, wie Daten aufbereitet worden sind (z. B. Transkriptions- und Anonymisierungsregeln), Datenmodifikationen (z. B. unterschiedliche Versionen der Daten), Dateienanzahl und in welcher Beziehung diese zueinanderstehen, spezifische Dateiinformationen (Dateititel, -formate, -größe) und schlussendlich eine Information darüber, ob es sich um sensible Daten handelt und wie diese genutzt werden können (vgl. 2018: 63).

Was bei den jeweiligen Datensätzen als Kontext gilt, kann dynamisch variieren, sodass dies in Form einer Offenheit in den Kontextualisierungsprozess miteinbezogen werden sollte. Dieser erfordert ebenfalls eine Abwägung zwischen der Sicherstellung des analytischen Potenzials und dem Umfang eingesetzter Ressourcen. Eine Herauslösung von narrativ

eingebetteten Forschungsdaten im Zuge einer Sekundärnutzung bedeutet zugleich, dass Forscher_innen teilweise ihre Interpretationshoheit abgeben müssen und potenzielle Nutzer_innen über relevantes Orientierungswissen zum adäquaten Verständnis der Daten verfügen sollten. Einschränkend ist jedoch zu beachten, dass implizites und intuitives Erfahrungswissen schwierig datenförmig darzustellen ist und dass die Aufzeichnungen nur einen geringen Teil der (körperlichen und emotionalen) Erlebnisse und Erinnerungen an die tatsächliche Begebenheit darstellen können (vgl. Breidenstein et al. 2013: 115; Imeri 2018b: 235). Die Daten sind keine objektive Abbildung von sozialer Realität, sondern haben bereits einen Interpretationsprozess durchlaufen und wurden auf eine wissenschaftliche Fragestellung zugeschnitten (vgl. Huber 2019a: 7) Meier zu Verl und Meyer merken kritisch an, dass wenn das Orientierungswissen nicht aufwendig „zur Sprache gebracht“ (2018: 87 f.) wird, die Daten größtenteils nicht nachvollziehbar bleiben. Einschränkend ist anzumerken, dass durchaus wissenschaftliche Publikationen (insbesondere Monografien) existieren, in denen dieses Wissen explizit thematisiert wird. Es müssen mitunter Wege gefunden werden, Selbstreflexionsprozesse und Emotionen der Forschenden zu beschreiben und potenziellen Nutzenden diese, durch dokumentierende Kontextinformationen, zur Verfügung zu stellen (vgl. Eberhard/Kraus 2018: 45 f.; Huber 2019a: 6). Hierzu gibt es bereits Lösungsansätze, umfangreiche Erfahrungswerte fehlen jedoch bislang.

Gleichzeitig wird den Sekundärforschenden, im Umgang mit den Forschungsdaten, deren Einordnung und Kritik abverlangt. Diese Herausforderungen sind jedoch für die Forschung kein grundsätzlich neues Problem — so muss sich bei derzeitiger geisteswissenschaftlicher Forschung auch ein Kontext erarbeitet und Quellenkritik geübt werden. Bishop und Kuula-Luumi (vgl. 2017: 10) weisen zudem darauf hin, dass selbst wenn relativ wenig Kontextinformationen verfügbar sind, die Sekundäranalyse mit positiven Ergebnissen durchgeführt werden kann. Wobei sie Datengebende trotzdem ermutigen, reichhaltige Projektbeschreibungen für die Nachnutzenden beizufügen.

Die Befragten wiesen einer sorgfältigen Datendokumentation als Nachnutzungsgrundlage eine hohe Bedeutung zu: Sie integrieren, wenn vorhanden, Transkripte, Forschungstagebücher, Selbstzeugnisse von Interviewten, Notizen, Karteikarten, Findbücher, Methodenberichte zur Erhebung der Daten oder auch Publikationen, die auf den jeweiligen Forschungsdaten basieren. Außerdem ergänzen sie Informationen darüber, in welchem Zusammenhang, von wem und auf welche Weise die Forschung realisiert worden ist, beispielsweise durch Projektanträge sowie Zwischen- und Abschlussberichte (vgl. TN3; TN4; TN5; TN6). Neben Materialien, die eher auf die

Makroebene abzielen, werden auch einzelnen Materialien Kontextinformationen beigelegt, beispielsweise: „Wie ist der Kontakt entstanden? [...] Wie ist das Interview abgelaufen [...]? Bis hin zu Nachüberlegungen, also welche ersten Überlegungen mag der Interviewer haben, was die Interpretation betrifft?“ (TN5: 68-71).

In Bezug auf den Forschungs- sowie den Kontextualisierungsprozess spielen Feldtagebücher und Feldnotizen eine wichtige Rolle, deren Nachnutzung, durch die persönliche Verwobenheit der Forschenden darin, umstritten ist (vgl. Imeri 2018b: 239 f.). Kritisch wurde von Forschendenseite angebracht, dass es nicht sinnvoll ist, diese zu archivieren, da sie zu persönlich sind und „ja keiner was mit anfangen (kann), so Zitat.“ (TN4: 20) Ein Ansatz, dieses Problem forschungsbegleitend produktiv zu lösen, wurde in einem anderen Interview angesprochen. Dort wurde eine Mischform, die sich zwischen Feldtagebuch und Datenmanagementplan bewegt, vorgeschlagen, die „beides irgendwie berücksichtigt und den Umgang mit den Daten einfach schon besser reflektiert. Weil das viel ersparen würde an späterem nachträglichem Aufbereiten des Materials und Kontextualisieren des Materials [...] für die Nachnutzung.“ (TN2: 552-55)

Ein Archiv, das sich derzeit viel mit bereits abgeschlossenen Forschungen beschäftigt, nutzt nichtstandardisierte forschungsbiographische Interviews, die zusammen mit eigenen Recherchen die Grundlage für die Erstellung von Kontextdaten darstellen. Dies wird als eine sehr sinnvolle Vorgehensweise, die es ermöglicht, wichtige Teile des Erfahrungswissens, die „Informationen im Kopf“ (TN3: 24) der Forschenden an mögliche Sekundärnutzende weiterzugeben, beschrieben. Im Anschluss werden einzelne Materialien gemeinsam mit den Primärforschenden durch weitere Informationen, die eine zeitlich-räumliche Einordnung möglich machen, ergänzt. (vgl. TN3: 22-28)

Die Kontextualisierung ist vielfach ein äußerst aufwendiger Prozess, weshalb ausreichend Ressourcen dafür zur Verfügung stehen müssen. So wurde in einem Interview konstatiert: „Ein Projekt, das keine Mittel für die Archivierung hat, dann ist es unheimlich schwer.“ (TN1: 70) Aus diesem Grund propagiert die Einrichtung, dass Forschende sich bereits in der Phase der Antragsstellung an sie wenden sollen, damit der Arbeitsaufwand abgeschätzt werden kann. Auf Grundlage dessen können dann, was vielfach geschieht, Anträge gestellt und die Arbeit der Forschenden an den Daten so finanziert werden (ebd.: 487-91). Trotzdem bleibt die Kontextualisierung von Beobachtungsprotokollen „einfach unglaublich viel Aufwand“, was bei einzelnen Ereignissen noch möglich ist, aber bei Forschung „über

Monate hinweg, Wochen hinweg [...] sind das so viele [...] Erlebnisse[n], Situationen, die in dieser Gesamtheit einfach nicht niedergelegt werden können.“ (TN2: 525-527)

Unkontrolliertes Videomaterial, insbesondere, wenn Dritte ins Bild kommen, erfordert einen hohen Kontextualisierungsaufwand, der oft unterschätzt wird. Da eine umfassende Anonymisierung ebenfalls sehr schwierig ist, ist hier vielfach die Nachnutzung auf eine gemeinsame Sekundäranalyse mit den Primärforschenden begrenzt, die im direkten Gespräch das Material umfassend kontextualisieren können (ebd.: 96-105). Kürzlich wurde ein Forschungsantrag zur Produktion von zu archivierendem Videomaterial gestellt, um erste Erfahrungswerte zu schaffen und rechtliche Fragestellungen, beispielsweise in Hinblick auf Einverständniserklärungen, zu eruieren (ebd.: 520-27).

Einschränkend wurde jedoch darauf hingewiesen, dass zahlreiche Kontextdaten nicht bei allen Forschungsdaten gleichermaßen vorhanden sind (vgl. TN5). So wird auch weniger gut erschlossenes oder noch zu digitalisierendes Material, beispielsweise Interviews ohne Transkriptionen, aufgenommen.

4.3.2) Auffindbarkeit der Forschungsdaten

Um nachgenutzt zu werden, ist es essenziell, dass Nutzende Kenntnis über die bereitgestellten Daten der Repositorien haben. Sinnvoll ist es, dass die Daten, bestenfalls in fachspezifischen Nachweissystemen, Fachportalen oder Bibliothekskatalogen auffindbar sind (vgl. Klump 2010: 112). Da es bislang für qualitative sozialwissenschaftliche Daten im deutschsprachigen Raum keine zentrale Einrichtung, wie das britische UK Data Archive gibt, ist die Wissensvermittlung, bezüglich der vorhandenen Ressourcen an Wissenschaftler_innen, von besonderer Bedeutung. Für die Nutzenden ist zu klären, wie viel Zeit sie beispielsweise in die Vorrecherche investieren können, anstatt nur die wichtigsten Punkte zu überfliegen (vgl. Corti/Fielding 2016: 10). In einem Interview wurde darauf hingewiesen, dass die Durchsuchbarkeit eine entsprechende Granularitätsebene und Datenbeschreibung erfordere, damit Forschende „möglichst schnell an möglichst genaue Informationen komme(n)“ (TN6: 503) und mit geringem Aufwand entscheiden können, ob diese für sie geeignet sind.

In Interviews wurde erwähnt, dass die Daten über Suchmaschinen wie Google auffindbar wären (vgl. TN1: 299-300; TN4: 248-49). Eine Infrastruktureinrichtung gab an, dass die eigenen Kataloge schon länger für Internetsuchen optimiert werden und kommentierte, dass die Aggregation in unterschiedliche Katalogsysteme durch Suchmaschinen „ein Stück weit [...] hinfällig geworden“ (TN6: 470-71) ist. Trotzdem besitzt die Einrichtung ein eigenes Katalogsystem, wo eine knappe Studienbeschreibung durchsuchbar ist (vgl. TN6: 77-83). Zugleich sind die Daten bei da|ra der Registrierungsagentur für sozialwissenschaftliche Daten, dem Datenkatalog des europäischen Verbunds CESSDA eingespeist und dort durchsuchbar (vgl. TN6: 456-75).

Eine andere Einrichtung beschrieb, dass auf die Daten aufmerksam gemacht würde, indem ein Datengeber, der aktiv forscht, archivierte Tonaufnahmen in seinen Artikeln verlinkt (vgl. TN3: 95-96). Die Anbindung an aktuelle Forschung war auch für ein weiteres Archiv bedeutsam, da insbesondere die Daten nach wie vor eine hohe Nutzung hätten, auf welche in rezipierten Publikationen verwiesen wird. Daten, die direkt ins Archiv integriert werden und nicht nochmals in anderen Kontexten Erwähnung finden, werden im Vergleich dazu deutlich seltener genutzt (vgl. TN5: 345-46).

Der Umgang mit expliziter Werbung ist unterschiedlich. Eine Einrichtung listet neue Datensätze auf ihrer Homepage ausdrücklich auf. Werden Workshops veranstaltet, werden diese über verschiedene Listen beworben, für die Betreuung von Social-Media-Kanälen fehlen derzeit jedoch die personellen Ressourcen (vgl. TN1: 308-17). Ein anderes Archiv wird zwar in Fachverzeichnissen gelistet, verzichtet aber darauf, selbst aktiv Werbung für die Materialien zu machen und sieht sich „in der Tradition des historischen Archivs“ (TN5: 227-28), wo es die „Aufgabe des Wissenschaftlers (ist) zu recherchieren.“ (ebd.: 231) Bei größeren personellen Ressourcen wäre dies jedoch denkbar, „also ich würde das schon für eine sinnvolle Aufgabe halten, sprengt aber im Moment völlig den Rahmen.“ (ebd.: 241-42) In einem anderen Interview wurde beschrieben, dass die Forschenden selbst für die Informationsweitergabe zuständig seien, das Archiv mitunter jedoch auch neues Material an daran interessierte Forschende weiterleitet. Bisher wurden, aus mangelnden zeitlichen Ressourcen, einzelne Datensätze nicht gesondert auf der Homepage ausgewiesen (vgl. TN4: 249-59). Eine Einrichtung strebt größere Sichtbarkeit an, die dann wiederum ein Argument für Forschende sei, Daten zu archivieren, um für ihre Projekte ein höheres Maß an Außendarstellung zu erreichen (vgl. TN3: 97).

Daneben bieten Einrichtungen Veranstaltungen an, um Interessierte zu informieren und miteinzubinden (vgl. TN1; TN3; TN6). Wenn Interesse vorhanden ist, können Workflows angeboten werden, mit denen die Forschenden praktisch arbeiten, samt Schulungsangeboten für technisches Equipment (vgl. TN4: 235-37). Diese Veranstaltungen haben oftmals den Effekt, Vorbehalte und Missverständnisse vonseiten der Forschenden abzubauen und ihnen Wissen über Angebote zu vermitteln. So wurde in einem Interview berichtet, dass viele Teilnehmende bezüglich Open Data davon ausgingen, dass ihre sensiblen Daten frei zugänglich sein sollten, was durch Hinweise auf hohe Sicherheitsstandards als Irrtum aufgelöst werden konnte (vgl. TN1: 454-65). Die Notwendigkeit einer Übersetzung von Begrifflichkeiten wie beispielsweise Metadaten wurde angemerkt, dessen produktive Anwendung und Potenzial vielen Forschenden nicht umfänglich bekannt sei (vgl. TN2: 562-67). Eine Einrichtung, die schon sehr lange Workshops zum Forschungsdatenmanagement anbietet, berichtete, dass es innerhalb der Ausrichtung der Veranstaltungen mittlerweile zu einer Verschiebung gekommen ist. Wurden anfänglich allgemeine Workshops, die Datenorganisation, -sicherung, -schutz sowie urheberrechtliche Aspekte thematisierten, angeboten, sind diese, da Universitäten vermehrt eigene Expert_innen dafür etablieren, weniger gefragt. Für die Zukunft sind deshalb sogenannte Train-the-trainer-courses geplant, die die Wissensweitergabe an diese Universitätsmitarbeiter_innen vorsehen. Zudem soll das Angebot spezialisiert werden, da insbesondere rechtliche Fragen hinsichtlich Anonymisierung und Urheberrecht „immer wieder Allen unter den Nägeln brennen“ (TN6: 216).

4.3.3) Kooperationen mit den Fachcommunities

Öffentlichkeit wird durch die Vorstellung bei Konferenzen oder durch Besuche an Universitäten geschaffen. Die Anbindung an die entsprechenden Fachcommunities ist, zwecks Etablierung von Kooperationen, wichtig. So kann eruiert werden, welche tatsächlichen Bedarfe, vonseiten der aktiv Forschenden an Infrastruktureinrichtungen bestehen (vgl. TN1: 253-59). Wissenschaftliche Entwicklungen sollen durch Angebote der Infrastruktureinrichtungen unterstützt und gefördert werden, Infrastrukturen könnten nicht bedarfsunabhängig vorausgehen und Angebote entwickeln (vgl. TN7: 180-90). Eine Interviewte beschrieb, dass Kooperationen dabei notwendig sind, weil sich beispielsweise ein dringend benötigter Thesaurus nur als gemeinsamer Standard über Institutionsgrenzen

hinweg schaffen lässt (vgl. TN4: 371-77). Eine strategische Zusammenarbeit über Fächergrenzen hinweg ist wünschenswert, da diese Einfluss auf Forschungspolicies nehmen könnte und spezifische Bedarfe, zum Beispiel, dass nicht das gesamte Material als open access zugänglich gemacht werden kann, artikulieren kann (vgl. TN4).

Bei der infrastrukturellen Vernetzung kommt der Zusammenarbeit mit den Fachinformationsdiensten eine wichtige Scharnierfunktion zu, um Öffentlichkeit zu generieren und Netzwerke zu erweitern (vgl. TN1: 360-62). Sie können durch ihre langjährige gute Vernetzung mit den Fachcommunities zwischen diesen und den Dateninfrastruktureinrichtungen, als Multiplikatorinnen, vermitteln. Sie können zugleich Rückmeldungen dahingehend geben, welche Bedarfe von deren Seite bestehen (vgl. Harbeck/Kaun 2019: 6, 8 f.).

Als ein Beispiel kann der FID Soziologie dienen, der zukünftig auf seinem Onlineportal Sociohub, neben einer bereits etablierten Literaturrecherchemöglichkeit, auch ein zentrales Nachweis- und Rechercheportal für publizierte Forschungsdaten etablieren will (vgl. Depping 2020: 5; Harbeck/Kaun 2019: 7). Daneben können sie Datennachweise in ihre eigenen Informationsinfrastrukturen, z. B. Rechercheplattformen, einbinden. Ihre Expertise wird bei der Etablierung der NFDI, als Zusammenschluss der Fachwissenschaft und Forschungsdateninfrastrukturen, nötig sein (ebd.: 3). Fachinformationsdienste sollten einerseits Ergebnisse der NFDI-Entwicklung an die eigene Forschungscommunity rückbinden und deren Informationslage verbessern und andererseits die spezifischen Fachbedarfe, die in der NFDI Eingang finden sollen, bei dieser zur Berücksichtigung anregen (vgl. Harbeck/Kaun 2019: 9 f.).

4.3.4) Datenauswahl und Arbeitsaufwand

Aus forschungsökonomischer Sicht sind, die zwei miteinander verbundenen Komplexe der Datenauswahl und die Frage nach dem Arbeitsaufwand, von besonderer Bedeutung: Zukünftig muss geklärt werden, welche Daten in welchem Aufbereitungsstand sinnvoll nutzbar sind. Da die Archivierung viele Ressourcen benötigt, ist eine allumfassende Datenarchivierung in der Zukunft nicht zielführend (vgl. Imeri 2018a: 74). Dem stimmen auch Knoblauch und Wilke zu: „Um die Schaffung teurer Datenfriedhöfe zu vermeiden, deren Daten von den Forschenden nicht genutzt werden, ist der durchgängige,

forschungspraktische und methodenpluralistische Einbezug der qualitativ Forschenden in die Entwicklung der Infrastruktur sicherzustellen.“ (2018: 54)

Smioski nennt Kriterien, um zu bestimmen, inwiefern Daten geeignet sind: Nach Möglichkeit sollten Projektdatensätze möglichst umfassend archiviert werden, beispielsweise Transkripte von Interviews, samt Aufnahmen und Protokollen. Eine Archivierung muss auf technischer Ebene möglich sein, Aufnahmen brauchen ein Mindestmaß an Aufnahmequalität, Datenformate müssen lesbar sein, dies gilt gleichermaßen für analoge wie digitale Daten. Vor der Archivierung müssen darüber hinaus urheberrechtliche und datenschutzrechtliche Fragen abgeklärt sein. Schlussendlich müssen die Daten nach den Grundsätzen guter wissenschaftlicher Praxis erfasst und dokumentiert werden (vgl. Smioski 2013: 9 f.).

Die DGfE nennt als Orientierungshilfe, in puncto Nachnutzungspotenzial, für die Primärforschenden bezüglich ihrer produzierten Daten: „die Datenqualität [...], die Verständlichkeit und Nachvollziehbarkeit [...], der Ausschöpfungsgrad [...], das Analysepotenzial [...], die Singularität der Daten [...], die Relevanz des Forschungsprojekts und seiner Ergebnisse [...] und die Kontextualisiertheit“ (2017: 6).

Die Qualität der Erhebung könnte beispielsweise auch durch Review-Verfahren beurteilt werden, wobei das data peer review als Qualitätssicherungsmaßnahme bislang selten angewandt wird (vgl. Huber 2019a: 15). Dies ist eine zusätzliche Aufgabe, bei der noch offen ist, wer diese zukünftig übernehmen könnte. Der dahingehende Umgang der Einrichtungen ist, bedingt durch Größe und Implementierungsgrad, unterschiedlich. Die Potenziale der Daten, während der Primäranalyse einzuschätzen, sei schwierig, was Umfangreiches, über die Publikation hinausgehendes, Material wünschenswert macht (vgl. TN2: 198-209). Die Auswahlkriterien für den Datenumfang obliegt bislang den Forschenden selbst, wobei die Einrichtung versucht, Anregungen zu geben, aber „die Entscheidungen können wir dann nicht vorwegnehmen, da sich das auch noch entwickelt“ (ebd.: 227-28). Ein Archiv begrenzt die Archivierung auf lebensgeschichtliche narrative Interviews, die offen für eigene subjektive Erzählungen sind, jedoch keine Expert_inneninterviews oder monothematischen Interviews (vgl. TN5: 149-70). Dies wird ergänzt durch andere (hand-)schriftliche Quellen, in denen sich Menschen an ihre Lebensgeschichte erinnern (ebd.: 45-48). Wobei „Ideal und Wirklichkeit“ (ebd.: 167) oftmals auseinanderlügen, denn die Daten entsprechen oft „nur sehr bedingt“ (ebd.: 158) den eigenen Anforderungen.

Eine andere Einrichtung nutzt eine ausformulierte Sammlungspolicy, die recht restriktionslos ist, solange die Datengebenden die Rechte an den Daten besitzen und sie „zur

Beantwortung von sozialwissenschaftlichen Fragen geeignet sind.“ (TN7: 203-04) Trotzdem existieren historisch gewachsene Schwerpunkte, bei denen bereits umfangreiche Bestände vorhanden sind. Obgleich es sich um eine große Infrastruktureinrichtung handelt, stehen trotzdem nicht ausreichend Ressourcen zur Verfügung, alle Daten im gleichen Maße aufzubereiten. Aus diesem Grund wird jetzt, nachdem man über „Jahre sehr opportunistisch rangegangen“ (ebd.: 228) sei, mit verschiedenen Kuratierungslevels gearbeitet, wobei bei allen Daten eine Eingangsprüfung auf Lesbarkeit und Anonymisierungsbedarf stattfindet.²¹ Das ressourcenarme Eingangslevel steht kostenlos zur Verfügung. Die Daten können dort selbständig über eine Software archiviert werden, wobei der physische Datenerhalt für mindestens 25 Jahre zugesichert wird (ebd.: 240-41). Darüber hinaus existieren weitere Services, die über die eigentliche Archivierung hinausgehen und preislich gestaffelt sind. Diese umfassen eine Qualitätsprüfung, kontinuierliche Formatmigration sowie eine ausführliche Dokumentation. Der Preis wird abhängig vom Datensatz und den gewünschten Leistungen berechnet.

Als arbeitsintensiv beschreibt eine Einrichtung das hohe Maß an Kommunikation mit den Forschenden selbst. So werden vielfach umfangreiche Vorgespräche geführt, in denen die Forschenden mitunter noch vom Nutzen und den Vorteilen der Archivierung überzeugt werden müssen und eruiert wird, was für Datenmaterial vorhanden ist, wie die Forschenden damit umgehen wollen und schließlich, was sich überhaupt eignet. Gründe für eine Nichtübernahme umfassen rechtlich-ethische Probleme, einen Mangel an Informationen über die Daten oder, dass der Arbeitsaufwand im Verhältnis zum Nutzen den Forschenden zu groß erscheint. Zudem sind manche kooperativ erzeugten Daten aufgrund der Einrichtungsstruktur problematisch, beispielsweise, „wo sehr stark mit Indigenen zusammengearbeitet wird und die administrative Rechte haben und vielleicht auch die Hoheit über das Löschen. Das Löschen ist derzeit nur sehr schwer möglich, weil das Repository [...] das lange Zeit nicht ermöglicht hat [...] Aber diese Nichtlösbarkeit und aber auch dieses Nichtmitarbeitsrecht, hat einige abgeschreckt.“ (TN4: 47-51)

Das Thema der Auswahl und Bewertung der Daten wurde in einem Interview als „ein ganz, ganz Zentrales. Um das sich bisher alle herumdrücken“ (TN7: 493-94) beschrieben. Man

²¹ Diese Einrichtung arbeitet hauptsächlich mit quantitativen Daten, sodass einschränkend festzuhalten ist, dass eine umfassende Übertragung dieser unterschiedlichen Kuratierungslevel auf den Umgang mit qualitativen Daten von anderen Archiven bislang noch nicht vorgesehen ist.

könne nicht „alle Daten für immer und auf höchstem Niveau [...] versorgen. Und es macht auch aus meiner Sicht überhaupt keinen Sinn.“ (ebd.: 495-97) Das nicht gänzlich absehbar ist, welchen Datenbedarf die Forschung in den kommenden Jahrzehnten hat, sei „ein gutes Argument dafür, möglichst viel aufzuheben.“ (TN7: 531-32) Jedoch „sich nicht zu entscheiden ist keine Option“ (ebd.: 554), auch wenn das möglicherweise bedeutet, dass einige zukünftig benötigte Datensätze nicht aufbewahrt werden. Selbst wenn alles technisch speicherbar wäre, würde vieles ohne ressourcenaufwendige tiefergehende Erschließung und Kontextualisierung nicht mehr auffindbar bzw. nutzbar sein (ebd.: 573-74). Zugespitzt wurde kritisch angemerkt, dass die „allermeisten Daten [...] keinen größeren Wert über das Forschungsprojekt hinaus, für das sie generiert“ (ebd.: 506-07) wurden, haben.

Die Entscheidung über die Archivierungswürdigkeit von Daten bedarf transparenter Mechanismen (vgl. TN7). Sie muss jedoch aus der Wissenschaftscommunity selbst kommen. Diese müsste Qualitäts- und Auswahlstandards entwickeln, damit Infrastruktureinrichtungen nicht „Gatekeeper für Ressourcen“ (ebd.: 321) werden oder einen „Mainstream prägen.“ (TN1: 170-71) Einrichtungen könnten lediglich den Diskurs darüber, was als archivierungswürdig gilt begleitend unterstützen, jedoch sei direkte Einflussnahme zu vermeiden. Die Aufgabe der Einrichtung sollte eher darin bestehen, zu beurteilen, inwiefern eine Studie ausreichend dokumentiert ist, damit auch andere damit arbeiten können. Da eine adäquate Archivierung nur mit Hilfe von eingeworbenen Ressourcen möglich ist, kann die Entscheidung darüber, durch den Mechanismus wissenschaftlicher Begutachtungsverfahren, an die Forschungsförderung zurückgegeben werden (vgl. TN7).

Als eine Möglichkeit schlägt eine etablierte Einrichtung eine Auswahl über einen transparenten Kostenmechanismus vor. Umfassende Services sollen nicht kostenfrei angeboten werden, sondern über Anträge finanziert werden. Im Zuge dessen kann auch überlegt werden, ob man ältere, dennoch bedeutsame Datensätze nachträglich durch Digitalisierung archiviert (vgl. TN1: 225-30).

Ein Archiv, das sich vor allem auf einen bestimmten Materialtyp fokussiert, betreibt bewusste Akquise auf Konferenzen oder fragt direkt bei Forschenden Daten an. Zunehmend würden Forschende, bedingt durch Anforderungen der Förderinstitutionen, schon zum Zeitpunkt der Projektbeantragung auf das Archiv zukommen und die Möglichkeit einer Archivierung erfragen. Dies sei „eine sehr komfortable Situation“ (TN5: 259), da dadurch Einflussnahme auf die (technische) Dokumentation bei der Datenerstellung möglich sei.

Die Frage der Auswahlkriterien spielt für Einrichtungen mit einem größeren Datenbestand eine andere Rolle als für Einrichtungen in der Etablierungsphase. Ein kleines

Archiv konzentriert sich zunächst auf das Material aus dem angrenzenden Universitätsinstitut und erweitert dieses sukzessive durch thematisch adäquate Begleitmaterialien, wie ergänzende Fotografien (vgl. TN4: 205-19).

Auf Seite der Forschenden muss sich zudem zeigen, inwiefern der aufwendige Prozess der Datendokumentation als Zusatzaufgabe wahrgenommen wird, die sie potenziell von ihrer eigenen Arbeit der Forschung abhält oder sie diese als produktive Ergänzung ihrer Arbeit ansehen.²² Dies wird auch davon abhängen, ob Datensammlungen „als originäre, in der wissenschaftlichen Reputationsökonomie verwertbare Leistung(en)“ (Imeri 2018b: 230) anerkannt werden und inwiefern dies als eine Motivation für Forschende dienen könnte (vgl. Queckbörner 2019: 66). Forschungsdaten haben bislang einen nicht annähernd gleichen Stellenwert wie andere Publikationen, sodass sich zeigen muss, inwiefern eine Aufwertung möglich ist, „denn Reputation lässt sich nicht verordnen“ (Hollstein/Strübing 2018b: 8). Ein Ansatz stellt ein von der ACUMEN-Gruppe entwickeltes Portfolio dar, das neben der Publikationszahl von wissenschaftlichen Artikeln und Monografien beispielsweise auch die Zahl der publizierten Datensätze und deren Downloadzahlen nutzt, um wissenschaftliche Leistung zu evaluieren (vgl. ACUMEN 2014: 36). Darüber hinaus bleibt abzuwarten, inwieweit die Daten von anderen Forscher_innen tatsächlich genutzt werden können (vgl. Huber 2019a: 16 f.).

Eine spätere Datenarchivierung erfordert zumeist einen höheren Grad an Datendokumentation, die deutlich einfacher ist, wenn sie bereits zu Beginn Teil der Forschung ist und in den Forschungsplanungsprozess integriert wird. Problematisch ist in diesem Zusammenhang, dass zu Forschungsbeginn oftmals noch nicht absehbar ist, inwiefern die Daten für eine spätere Archivierung geeignet sein werden. Trotzdem wird vielfach ein Datenmanagementplan angestrebt, in dem festgehalten werden soll, welche Daten für eine Sekundärnutzung relevant sein können und in welcher Form sie erfasst werden (vgl. Smioski 2013: 10). In vielen Einrichtungen ist deshalb eine

²² Vgl. hierzu: „Zudem muss der Aufwand für die Aufbereitung der Daten bewältigbar bleiben und ein sinnvolles Verhältnis zum Ertrag gewahrt werden. Das Datenmanagement ist daher eine Gratwanderung zwischen notwendiger Standardisierung von Strukturen, Prozessen und Workflows einerseits, ohne die kein effizientes und professionelles Datenmanagement auskommt, und den Anforderungen an Heterogenität, Kontextualisierung und Qualität der Forschungsdaten andererseits, ohne die eine inhaltlich ertragreiche Re-Interpretation nicht möglich ist.“ (Dunkel/Hanekop 2018: 40)

forschungsbegleitende Archivierung entweder schon vorhanden oder soll etabliert werden (vgl. TN1; TN4; TN5).

So beschrieb eine Einrichtung die Vorteile, wenn bereits zu Forschungsbeginn eine spätere Archivierung anvisiert wird und der Forschungsprozess durch das (planerische) Wissen der Forschungseinrichtung unterstützt wird, dass „damit nicht zum Schluss eine Gedächtnisleistung ansteht, die ein Berg von Arbeit ist, die dann häufig gar nicht mehr gemacht werden kann“ (TN1: 111-13). Zugleich ergeben sich für die Primärforschung selbst Vorteile, denn „durch diese gute und stringente Dokumentation habe ich natürlich eine ganz andere Qualität in meinem Primärforschungsprojekt. [...] Ich bewahre mein Wissen im Projekt, falls irgendwelche Personalwechsel sind.“ (ebd.: 118-23)

Die Archivierung sollte, insbesondere, weil die Aufbereitung des Materials sehr arbeitsintensiv ist, nach Möglichkeit zeitlich unbegrenzt sein (vgl. Imeri et al. 2019: 17), da vielfach vermutet wird, dass der Wert der (ethnografischen) Daten für die Forschung, je älter sie sind, anwachsen könnte (vgl. Imeri 2018b: 238), wozu eine nachhaltige Finanzierung notwendig ist, die oft Thema der Interviews war (s. u.) (vgl. TN1; TN3; TN4; TN6).

4.3.5) Zwischenfazit praktisch-technische Probleme im Umgang mit qualitativen sozialwissenschaftlichen Forschungsdaten

Die Vielfältigkeit der angewandten Forschungsmethoden hat zur Folge, dass die produzierten qualitativen Forschungsdaten ein geringes Maß an Standardisierung (z. B. Software, Speicherformate) aufweisen. Dies schafft verstärkten Arbeitsaufwand durch die Entwicklung von individuelleren Workflows und Qualitätsstandards. Eine generische Herangehensweise stellt der Rückgriff auf Datenmanagementpläne dar, der jedoch von Forschenden mitunter kritisch gesehen worden ist (vgl. TN2: 618). Erschwerend kommt hinzu, dass sich durch die Offenheit des Forschungsprozesses das Archivierungs- und Nachnutzungspotenzial schwierig im Vorhinein genau bestimmen lässt (ebd.: 616-24).

Die Auffindbarkeit von Forschungsdaten macht eine fundierte Erschließung notwendig. Hierbei wird versucht eine Balance zwischen generisch-austauschbaren und spezifisch-kontextualisierten Beschreibungen mittels Metadaten zu erreichen. Die interviewten Einrichtungen nutzen in der Praxis deshalb beispielsweise Metadatenschemata, die Pflichtfelder enthalten, gleichzeitig aber auch durch Freitextfelder an Spezifika

anpassbar sind (ebd.: 624-627). Problematisch sei in der Vergangenheit gewesen, dass die Metadaten mitunter überladen wurden (vgl. TN3: 35) und jederzeit einsehbar gewesen seien (ebd.: 37). Neben der händischen Auszeichnung mit Metadaten, versucht ein Archiv auf automatisierte Spracherkennung zurückzugreifen, insbesondere da das gesamte Material aufgrund des Umfangs und Ressourcenmangels nicht zeitnah umfassend erschlossen werden kann (vgl. TN5: 194-207).

Die Erschließung ist darüber hinaus abhängig von kontrollierten Vokabularen, die eine größere Interoperabilität und verbesserte Auffindbarkeit ermöglichen. Während eine Institution zwecks Inhaltserschließung auf die deutschsprachige Version des European Language Social Science Thesaurus (ELSST) zurückgreift (vgl. TN6: 186), wurde mehrfach der Mangel an einem adäquaten Vokabular in Hinblick auf qualitative Daten problematisiert. Dies sei nicht existent oder unbrauchbar (vgl. TN1: 283-84; TN3: 114). Ein Ausbau ist deshalb notwendig z. B. muss die Gemeinsame Normdatei (GND) entsprechend rezenter Fachdebatten überarbeitet und ausgebaut werden.

Eine dauerhafte Datenreferenzierbarkeit benötigt den Rückgriff auf persistente Identifikatoren, wobei die interviewten Einrichtungen zu diesem Zweck auf Digital Object Identifier (DOI) zurückgreifen (vgl. TN4: 248; TN6: 458-60).

Die spätere Nutzung qualitativer sozialwissenschaftlicher Daten macht vielfach ein hohes Maß an zusätzlicher Datendokumentation jenseits kompakter Metadaten nötig. Diese Kontextualisierung ist häufig ein recht zeitaufwändiger und ressourcenintensiver Prozess. Sie umfasst Informationen über die Forschungszusammenhänge z. B. Projektanträge oder Abschlussberichte, aber auch Material wie beispielsweise Transkripte, Forschungstagebücher, Findbücher, Methodenberichte bis hin zu nichtstandardisierten forschungsbiographischen Interviews (vgl. TN3; TN4; TN5; TN6).

Bei Feldtagebüchern und Notizen erschwert die Involviertheit der Forschenden darin deren Weitergabe, da sie einerseits zu persönlich seien und andererseits „keiner was mit anfangen (kann), so Zitat.“ (TN4: 20) In Reaktion auf diese Einwände schlägt eine Einrichtung eine Mischform vor, die versucht Feldtagebuch und Datenmanagementplan miteinander zu verbinden (vgl. TN2: 552-55).

Die Auffindbarkeit der Forschungsdaten wird über Suchmaschinen sichergestellt (vgl. TN1: 299-300; TN4: 248-49). Eine Einrichtung hob hervor, dass die eigenen Kataloge für die Suchmaschinenrecherche bereits optimiert seien (vgl. TN6: 470-71). Vielfach stellt auch eine Rückbindung an (gegenwärtige) Forschung beispielsweise durch allgemeine Hinweise auf Datenbestände in Publikationen (vgl. TN5: 345-46) oder durch unmittelbare

Verlinkung der Forschungsdaten (vgl. TN3: 95-96) eine Möglichkeit zur Sichtbarmachung von Material dar.

Die Kooperation mit den Fachcommunities, mitunter vermittelt über Fachinformationsdienste, ist ebenfalls bedeutsam. Einerseits kann durch die Vorstellung bei Konferenzen oder bei Universitäten das eigene Angebot bekannter gemacht werden und andererseits auf diese Weise Bedarfe der Forschenden eruiert werden, auf welche die Infrastruktureinrichtung wiederum reagieren kann (vgl. TN1: 253-59).

Der benötigte Ressourcenaufwand für die Aufbereitung und dauerhafte Archivierung ist recht hoch, sodass eine umfassende Archivierung großer Materialmengen zukünftig nicht realisiert werden kann. Vielmehr muss zukünftig präziser bestimmt werden, welche Daten zur Archivierung ausgewählt und mit welchem Arbeitsaufwand sie weiter aufbereitet werden sollen, um sinnvoll nutzbar zu sein. Dies könnte über transparente Kostenmechanismen und Review-Verfahren beispielsweise, indem über die notwendigen Ressourcen durch die Institutionen der Forschungsförderung selbst entschieden wird, geschehen. Bislang orientiert sich eine Einrichtung bei der Auswahl an der Art der geführten Interviews (vgl. TN5: 149-70), eine andere verfügt über eine ausformulierte Sammlungspolicy, die jedoch recht breit gefasst ist (TN7: 203-04). Auf Seiten der Forschenden muss sich schließlich herausstellen, inwieweit die aufwändige Datendokumentation als Zusatzaufgabe angesehen wird oder ob jene diese als konstruktive Bereicherung ihrer Tätigkeit betrachten.

4.4) Nachnutzung qualitativer sozialwissenschaftlicher Forschungsdaten

Die Wiederverwendung und Untersuchung von Daten früherer Forschungsprojekte bietet zunehmend die Gelegenheit, sowohl für methodische als auch für inhaltliche Zwecke, neue Erkenntnisse zu generieren sowie neue Überprüfungsansätze zu etablieren. Corti et al. beschreiben, dass der Nachnutzungsansatz in den letzten Jahren zugenommen hat, wobei Faktoren wie die Open-Data-Bewegung, die Politik der Forschungsförderungsorganisationen, die die Forschungstransparenz und die gemeinsame Nutzung von Daten unterstützen und Forschende, die die Vorteile der gemeinsamen Nutzung aller möglichen Ressourcen sehen, eine Rolle gespielt haben (vgl. 2016: 2).

Bezüglich der Offenlegung von Betrugsfällen durch die Veröffentlichung von Forschungsdaten, wenden Pels et al. jedoch kritisch ein, dass die jüngsten Fälle in den Niederlanden, in denen wissenschaftlicher Betrug vermutet wurde, nicht die Hypothese stützen, dass die Veröffentlichung von Rohdaten dazu beiträgt, Betrug leichter aufzudecken oder zu verhindern. Standardverfahren zur Überprüfung von wissenschaftlicher Integrität beispielsweise durch Expert_innenbegutachtung hätten ausgereicht, um Täuschungen auszuschließen (vgl. 2018b: 400).

Die Organisationen der Forschungsförderung versprechen sich aus forschungsökonomischer Sicht, dass effizienter gearbeitet werden kann, indem gleichartige Daten nicht erneut gewonnen, sondern mehrmalig verwendet werden sollen (vgl. Imeri 2018b: 216 f.). Formen der Nachnutzung umfassen die Sekundäranalyse, dabei werden die erhobenen Daten im Rahmen neuer Fragen erneut untersucht oder die Daten als Erweiterung in den Datenbestand neuer Erhebungen integriert (vgl. DGfE 2017: 3).²³

Einige Probleme der Wiederverwendung von Daten werden von Corti beschrieben und beziehen sich auf: Bedenken hinsichtlich der ethischen Wiederverwendung von Daten; dass nicht genügend Daten zur Analyse verfügbar sind; Unkenntnis der Methode der Sekundäranalyse; Mangel an kontextbezogenen Informationen; verzerrte Auswahl und selektive Stichproben bei der Sekundäranalyse (vgl. 2018: 20).

Im Gegensatz zur Arbeit mit sozialwissenschaftlichen quantitativen Daten, wo eigenständige Datenproduktion „nicht das Maß aller Dinge“ (TN7: 368-69) ist und Sekundäranalysen vielfach stattfinden, ist dies in vielen Disziplinen, die mit qualitativen Daten arbeiten weit weniger, u. a. bedingt durch die oft enge Verzahnung von Datenerhebung und -auswertung, verbreitet. Dies sollte von den Datenzentren berücksichtigt werden (vgl. Sahle/Kronenwett 2013: 82).

So kommentiert Smioski: „Die Archivierung und Sekundäranalyse quantitativer Daten sind fest in den Sozialwissenschaften verankert. Das UK Data Archive (UKDA), 1967 gegründet, feierte 2007 sein 40-jähriges Jubiläum und CESSDA, das Council of European Social Science Data Archives, besteht seit den 1970er-Jahren. Auch die Re-Analyse

²³ Für eine Auswahl von Sekundäranalysebeispielen auf der Basis von Interviewdaten sowie Studien, die gezielt die Potenziale, Herausforderungen und Probleme von Sekundäranalysen qualitativer Daten beleuchten, siehe: Hollstein/Strübing (2018b: 6). Sie konstatieren den großen Bedarf nach „reflektierten Praxisbeispielen und beispielhaften Studien („best practice“), die zeigen, wie der methodologisch angemessene Gebrauch dieses Materials, auch in Kombination mit Primärdaten einer gemischt operierenden Studie, aussehen kann.“ (ebd.) Bishop und Kuula-Luumi (2017: 1 f.) weisen darauf hin, dass dieser Forschungsansatz der qualitativen Sekundäranalyse immer häufiger in Methoden-Lehrbüchern dargestellt wird und somit zu einer etablierten und anerkannten Forschungsmethode geworden ist. Für konkrete Beispiele siehe ebd.

nationaler und internationaler Umfrageprogramme ist aus der Forschung nicht wegzudenken. [...] Die Debatte um die Archivierung und Bereitstellung qualitativer Daten ist wesentlich jünger und hat erst Mitte der 1990er-Jahre begonnen. Die Gründung des britischen Archivs ESDS Qualidata im Jahr 1994, als integraler Bestandteil des UK Data Archives, stellt in dieser Agenda einen wichtigen Meilenstein dar. [...] Anfangs lag das Augenmerk auf der Archivierung klassischer Studien, die für eine weitere Nutzung bewahrt werden sollten. [...] Die Archivierung neuerer Studien wurde jedoch wichtiger, als das Economic and Social Research Council (ESRC) im Jahr 2000 eine Datenpolitik verlautbarte, die eine Archivierung von Forschungsdaten nach Projektende vorschreibt.“ (ebd. 2013: 2)

Ein großer Teil der Sekundäranalysen ist bisher anhand von Interviewprotokollen erfolgt, da die Tonaufnahmen in vielen Fällen nicht archiviert wurden oder den technisch-datenschutzrechtlichen Voraussetzungen nicht entsprechen (vgl. Huber 2019a: 7 f.). Daten werden zumeist neu und an die jeweilige Forschungsfrage angepasst generiert (vgl. Hollstein/Strübing 2018b: 2 f.).

Wissenschaftliche Karrieren basieren oftmals auf der Durchführung eigener Feldforschung/Datenerhebung (Imeri 2018b: 231). Unklar ist, in welchem Maße sich auf Datenebene eine „Kultur der Zweitverwertung“ (Sahle/Kronenwett 2013: 82) etablieren wird. Trotz vieler offener ethischer und erkenntnistheoretischer Hürden, beispielsweise hinsichtlich der Nachvollziehbarkeit durch Außenstehende, bietet die Sekundäranalyse, via Kontextualisierung und Quellenkritik, sowie durch die vorhandene Informationsdichte die Chance einer Vervielfältigung von Interpretationen und Reflexionen, die zu neuen Erkenntnissen führen kann (vgl. Huber 2019a 12, 20). Dies gilt insbesondere, weil, wie in einem Interview hervorgehoben, das Material stets nur teils ausgewertet ist und die Möglichkeit für weitergehende Forschungsanschlüsse bietet (vgl. TN3: 66-68). Die Analysen können zudem durch die Offenlegung methodischer oder ethischer Mängel als Kritikgrundlage genutzt werden.

4.4.1) Konkreter Ablauf für die Nachnutzung

Die Einrichtungen sehen unterschiedliche Herangehensweisen zum Zwecke der Nachnutzung vor: Generell gilt, dass das meiste Material zugangsbeschränkt ist und deshalb nicht unmittelbar von außen darauf zugegriffen werden kann, sondern zuerst mit der Einrichtung Kontakt aufgenommen werden muss. Die Daten sind somit nicht im Verständnis

von „open data“ absolut offen zugänglich, sie weisen vielmehr eine, wie oben im Rückgriff auf die FAIR-Prinzipien beschrieben, prinzipielle Zugänglichkeit auf (vgl. Wilkinson et al. 2016: 3). Wenn bereits bekannt ist, welche Daten benötigt werden, erhalten die Forschenden eine Nutzungsvereinbarung, die Angaben zum Zweck, der Dauer und der Finanzierung der Forschung enthält und danach den Datenzugang (vgl. TN1; TN5).

Falls interessierte Forschende nicht wissen, welche Daten für sie geeignet sind, unterstützen Einrichtungen beratend bei der Recherche. Eine Interviewte berichtete, dass derzeit ein Daten-managementsystem zwecks Recherchemöglichkeiten entwickelt wird. Damit wird neben der Recherche von außen, via öffentlichen Metadaten, die studienspezifische Fallrecherche mittels Mikrometadaten (z. B. Alter, Geschlecht) im Safe-Room möglich sein (vgl. TN1: 388-98).

In einem anderen Archiv erfolgen Suchen über Metadaten, die Angaben zu Alter, Geschlecht oder Ort des Interviews machen. Sind darüber hinaus inhaltliche Kriterien (z. B. Lebenserfahrungen) von Bedeutung, werden mit den Forschenden gemeinsam sinnvolle Suchbegriffe generiert. Als Ergebnis erhalten die Forschenden Transkriptausschnitte, auf deren Grundlage sie entscheiden können, ob sie das Interview vor Ort benutzen wollen. Bis auf eine Projektliste, können die Forschenden nicht von außerhalb eine Recherche durchführen, wobei derzeit versucht wird, ein Online-Archiv aufzubauen (vgl. TN5: 268-70).

In einem Archiv in der Aufbauphase gab es bislang noch keine externen Nachnutzungsanfragen. Potenziell würde bei Anfragen eine Einführung gegeben und bei Bedarf auch ein Kontakt zu Primärforschenden hergestellt (vgl. TN4: 284-88).

Nur die eher quantitativ arbeitende Einrichtung besitzt ein automatisiertes Downloadsystem, wo Nutzende nach einer Registrierung im System nach nicht-sensiblen Daten recherchieren und diese selbstständig herunterladen können. Falls nichts Passendes gefunden wird, können Nutzende Recherchenachfragen stellen und bei technischen Problemen wie Fehlermeldungen um Hilfe bitten (vgl. TN7: 66-74).

Die Datennachnutzung wurde, in mehreren Interviews, als nicht nur auf ein Forschungsprojekt beschränkt beschrieben (vgl. TN1; TN5; TN6). In einem Archiv finden häufig Quernutzungen statt, was bedeutet, dass quer zu den in sich geschlossenen Projekten nach forschungsrelevanten Themen gesucht wird und so passende Informationen aus mehreren Datensätzen zusammen betrachtet werden (vgl. TN5: 309-13). Zudem wurde

beschrieben, dass Daten an Forschende weitergegeben wurden, die sich mit Methodenforschung beschäftigen, beispielsweise welche Eingangsfragen im Anschluss welche Erzählungen erzeugen (vgl. TN5: 318-23).

Wissenschaftler_innen nutzten die Daten, zwecks Forschungsvorbereitung und Fragengenerierung, beispielsweise dahingehend, Forschungslücken zu ermitteln. Daneben werden die Daten als Ergänzung bei Promotionsprojekten oder Masterarbeiten genutzt. Da eigentlich nur Wissenschaftler_innen die Einrichtung nutzen dürfen, wird bei Masterstudierenden die Datenbereitstellung stärker geprüft. Ein besonderes Nutzungspotenzial liegt in der Kombination quantitativer und qualitativer Daten sowie von archivierten Materialien mit selbstständig erhobenen, entweder im Langzeitvergleich oder zwecks Gegenüberstellung (vgl. TN1). Eine Einrichtung beschrieb als Nachnutzungsvariante Anfragen, zu Vergleichszwecken, bezüglich älterer vorhandener Daten, bei denen ein Kontakt zu den damals Beforschten hergestellt werden soll (ebd.: 415-18).

In einem Projekt ist man mit der Aufbereitung des Materials noch nicht so weit vorangeschritten, dass eine Bereitstellung möglich wäre. Angenommen wird jedoch, dass das Material eher eine Inspiration darstellt und nicht als ausschließliche Grundlage für eine Forschung genutzt wird. Dies ist auch durch die Materialanforderungen der Nachnutzenden bedingt, „man vertraut [...] sich doch dann wahrscheinlich mehr in der Forschungssituation als Material, das andere produziert haben.“ (TN2: 465-66)

Eine potenzielle Nachnutzung muss nicht notwendigerweise die Datengebenden und -nutzenden trennen. In einem Interview wurde von einem produktiven Workshop berichtet, der beide Gruppen zusammenbrachte. Ein positiver Effekt umfasste für viele Primärforschende dabei die Erkenntnis: „Das ist toll, wofür die Daten jetzt noch genutzt werden können. Ich wäre da nie draufgekommen [...] Die wurden für völlig unterschiedliche Themen genutzt und das fanden die Primärforschenden sehr beeindruckend.“ (TN1: 499-503) Solche Workshops sind sinnvoll, um nach wie vor virulente Vorbehalte und offene Fragen zu klären, sowie die Potenziale einer (kollaborativen) Sekundäranalyse deutlich zu machen. Insbesondere im ethnologischen Kontext könnten zukünftig für Ergänzungen offene Datenbanken/-archive Chancen bieten, Beforschte, wenn gewünscht, verstärkt in die Datenauswertung miteinzubeziehen (vgl. Huber 2019a: 13).

Daneben gibt es die Möglichkeit, eine Reanalyse durchzuführen, in der archiviertes Forschungsmaterial genutzt wird, um sich kritisch mit vorhandenen Forschungsergebnissen auseinanderzusetzen ohne, dass Daten gänzlich neu erhoben werden müssen (vgl. DGfE 2017: 3). Replikationsstudien zielen auf das exakte Reproduzieren der wissenschaftlichen Ergebnisse ab, wobei dies im Umgang mit qualitativen Daten jedoch weder möglich noch nötig ist. Ergebnisse qualitativer Forschung lassen sich nicht unmittelbar replizieren, sodass die mit Replikationsstudien verbundenen Hoffnungen nach erhöhter Glaubwürdigkeit und Forschungstransparenz sich nicht umsetzen lassen (vgl. von Unger 2018: 93 f.). Das methodische Vorgehen sieht notwendigerweise eine interpretative Erschließung des Materials vor, sodass die deckungsgleiche Übernahme von naturwissenschaftlichen Qualitätskriterien für den Forschungsprozess nicht zielführend ist (vgl. Bogner et al. 2014: 92 f.). Trotzdem existieren funktionierende Verfahren, um die wissenschaftliche Qualität zu sichern, beispielsweise, indem der Erhebungs- und Auswertungsprozess transparent gemacht wird oder durch die gemeinsame Interpretation der erhobenen Daten in Forschungswerkstätten (vgl. DGfE 2017: 3 f.; dgV 2018: 2), die sich zukünftig durch die Archivierung der Forschungsdaten durchaus unterstützen ließen.

In den geführten Gesprächen spielten diese auf Verifizierung und Qualitätskontrolle abzielenden Ansätze eine untergeordnete Rolle, wobei in einem Interview durchaus ein wachsendes Interesse, bezüglich der Bereitstellung der erhobenen Forschungsdaten zwecks Überprüfbarkeit der Ergebnisse beschrieben, wurde (vgl. TN5: 42-45).

Inwiefern Nachnutzung und Sekundäranalysen in Kombination mit neuen digitalen Methoden zukünftig die Forschungs- und Publikationspraxis beeinflussen werden, bleibt abzuwarten. Derzeit existieren erweiterte Publikationen (enhanced publications), was die Verknüpfung einer digitalen Publikation mit Daten wie Quelldokumenten, Audio- oder Videosegmenten, Grafiken, Modellen umfasst (vgl. Corti/Fielding 2016: 10). In einem Interview wurde berichtet, dass dies durch Verlinkung wissenschaftlicher Artikel mit Audioaufnahmen, die im Repository gespeichert wurden, geschieht (vgl. TN3: 95-96).

Die Sekundärnutzung von Daten kann auch Teil der Lehre sein, wobei Bishop und Kuula-Luumi beschreiben, dass die Mehrzahl der Lehrprojekte die Vermittlung qualitativer Forschungsmethoden betraf (vgl. 2017: 4). Für die Lehre waren insbesondere Sammlungen mit unterschiedlichen Datentypen, Daten, die für aktuelle politische Debatten wichtig sind, und Themen, die für junge Menschen von Interesse sind, von Relevanz (vgl. Bishop/Kuula-Luumi 2017: 12). Ihre Ergebnisse zeigen zudem, dass drei Fünftel der Datensammlungen

des UK Data Service für Lehren und Lernen genutzt werden und dass diese Nutzung zunimmt (ebd.: 13). Das aktive Bewerben einer Studie, entweder durch ein bestimmtes Fakultätsmitglied oder durch den UK Data Service, erhöht die Datennutzung erheblich (ebd.: 6).

Die Herausforderung bei der Entwicklung ansprechender Lehrmittel liegt im Zeitaufwand und in den Kosten für deren Herstellung, dabei kann oft eine Strategie der Koproduktion mit Datengebenden eingesetzt werden. Obwohl ein Einsatz von Ressourcen erforderlich ist, gibt es zwei offensichtliche Vorteile für Archive: Die Nutzungszahlen der Daten werden steigen, und eine nächste Generation von Forschenden wird lernen, die Wiederverwendung qualitativer sozialwissenschaftlicher Daten als Routine und normalen Teil ihrer Methodenausbildung zu sehen.

Viele der Einrichtungen beschrieben, dass ihre gesammelten Daten für die Lehre und (Methoden-)Ausbildung eingesetzt werden (vgl. TN5: 318-19), mitunter gibt es gar die meisten Nachfragen für die Lehre (vgl. TN1: 421-22). Dies reicht von Datengebenden, die eigenes aufbereitetes Material regelmäßig als konkrete Beispiele für Lehrveranstaltungen nutzen (vgl. TN3: 64-65; TN4: 327-28), hin zu externen Anfragen zur Datenbereitstellung. Zudem werden sogenannte knowledge packages erstellt, für die sich Lehrende bei der Einrichtung melden können, um nochmals speziell anonymisierte Transkripte samt Kontextualisierungsdokumenten nutzen zu können (vgl. TN1: 526-33). Wobei manche Themen, beispielsweise sexualisierte Gewalt, als zu sensibel für die Weitergabe an Studierende angesehen werden (ebd.: 425-27). Für audio-visuelle Materialien ist erst noch eine datenschutzrechtliche Klärung nötig und es fehlt vielen Universitäten an technischer Infrastruktur zur adäquaten Nutzung (ebd.: 537-44).

Eine Institution will bei Daten, die bereits in die Lehre gegeben wurden, eruieren, wie die Daten verwendet wurden und wie Studierende darauf reagiert hätten (ebd.: 518-20). Die Sekundärnutzung im Zuge der Lehre sei „ein starkes Argument“ (ebd.: 512) für die Einrichtung. Das forschungsnahe Lernen, an archivierten Forschungsdaten, würde ein Bewusstsein für deren Wert schaffen und eine Normalisierung der Arbeit mit den Daten ermöglichen: „Ich glaube diesen Schritt müssen wir in den nächsten Jahren schaffen.“ (ebd.: 514)

Eine Einrichtung mit größerem Fokus auf quantitative Daten gab an, dass Studierende „schon seit langem tatsächlich trainiert (werden) mit solchen Daten eigene Fragestellungen zu entwickeln und zu beantworten“ (TN7: 358-59). In vielen

Fachdisziplinen, die mit qualitativen Daten arbeiten, ist dies noch ein neueres, weiter zu etablierendes Vorgehen.

In einem Interview wurde zudem beschrieben, dass durch Generalisierung von Feldprotokollen und Feldnotizen sogenannte Case Studies entwickelt werden sollen, die Forschungssituationen beschreiben und als methodische Vorbereitung für Forschungsvorhaben dienen sollen (vgl. TN2: 477-84). Mit Hilfe gut beschriebener methodischer Reflexionen ließe sich potenziell viel lernen, was beispielsweise Feldzugänge zu schwierig zu erreichenden Gruppen angeht. Insbesondere Forschungen, bei denen der Verlauf nicht ideal gewesen sei, könnten besonders erkenntnisfördernd sein. Voraussetzung dafür ist jedoch, dass die Forschenden durch die Datenoffenlegung keine Nachteile im Wissenschaftsbetrieb davontragen (ebd.: 494-501).

4.4.2) Zwischenfazit Nachnutzung qualitativer sozialwissenschaftlicher Forschungsdaten

Innerhalb der Interviews wurde deutlich, dass derzeit bereits Formen der Nachnutzung praktiziert werden. Bei manchen Projekten ist die Datenaufbereitung allerdings noch nicht weit genug entwickelt, um die Daten schon zur Verfügung zu stellen.

Damit eine Nachnutzung realisiert werden kann, müssen Nutzende Zugang zu den Daten erhalten. Das meiste Material der Einrichtungen kann jedoch nicht unmittelbar verwendet werden, da es Zugangsbeschränkungen unterliegt. Eine Anmeldung und das Ausfüllen einer Nutzungsvereinbarung, die das Ziel und das Vorgehen innerhalb der Forschung darstellen, sind notwendig, um Zugang zu den archivierten Daten zu bekommen (vgl. TN1; TN5).

Wenn unklar ist, welche Daten von Relevanz für die Forschung sein könnten oder die vorhandenen Metadaten zu wenig aussagekräftig sind, unterstützen die Einrichtungen zudem mitunter bei der Recherche. Dies geschieht, indem sie eine studienspezifische Fallrecherche durch Mikrometadaten (z. B. Alter, Geschlecht) im Safe-Room zu Verfügung stellen (vgl. TN1: 388-98) oder gemeinsam mit den Forschenden Suchtermini generieren, die im Anschluss zu Recherche in Transkripten verwendet werden können. Ein anderes Archiv bietet den Nutzenden zusätzlich die Möglichkeit Rechercheanfragen zu stellen, wenn deren eigene Suche nicht erfolgreich gewesen ist (vgl. TN7: 66-74).

Formen der Nachnutzung umfassen die Sekundäranalyse, die archivierte Daten nutzt, die oft nur zum Teil ausgewertet sind, um auf deren Grundlage neue Fragen zu bearbeiten.

Außerdem werden die Daten als eine Ergänzung des Materials, das in einer aktuellen Forschung erhoben worden ist, verwendet. Ansätze, die Forschungsergebnisse verifizieren bzw. deren Qualität kontrollieren wollen, stellen keine zentrale Nutzungsformen innerhalb der Einrichtungen dar.

Die bereits praktizierten Nachnutzungen können über ein einzelnes Forschungsprojekt hinausgehen (vgl. TN1; TN5; TN6). Sie umfassen Quernutzungen, wobei in Datensätzen unterschiedlicher Forschungsprojekte nach Informationen gesucht wird, die für die eigene Forschung relevant sind. Quantitative und qualitative Daten können kombiniert werden und die archivierten Daten für einen Langzeitvergleich verwendet werden. Daneben werden die Daten für die Methodenforschung oder dafür genutzt Forschungslücken zu eruieren, an die die eigene Forschung anschließen kann. Datengebende und –nutzende können im Zuge der Nachnutzung durchaus produktiv kollaborativ zusammenarbeiten, wie eine Einrichtung anhand eines von ihr, für beide Gruppen gemeinsam, organisierten Workshops beschrieb (vgl. TN1).

Darüber hinaus werden die Daten oftmals in Ausbildungskontexten bzw. der Lehre eingesetzt, wobei mitunter die stärkste Nachfrage aus diesen Bereichen kommt (vgl. TN1; TN3; TN4). Hierbei wird zum Beispiel auf zweckspezifische Transkriptionen sowie weitere Dokumente zur Kontextualisierung zurückgegriffen (vgl. TN1: 526-33). Die gleiche Einrichtung wies in diesem Zusammenhang darauf hin, dass zur Nutzung von audiovisuellem Material noch datenschutzrechtliche Fragestellungen zu klären sind und eine angemessene Nutzung durch den Mangel an technischer Infrastruktur an vielen Universitäten bislang nicht möglich sei (ebd.: 537-44). Die Sekundärnutzung im Bereich der Lehre fungiert als „starkes Argument“ (ebd.: 512) für den Fortbestand der Einrichtung, der Rückgriff auf Daten innerhalb der Lehre normalisiere zugleich deren Nutzung und mache den Wert der Materialarchivierung deutlich.

4.5) Herausforderungen und Hindernisse in Bezug auf das Management qualitativer sozialwissenschaftlicher Forschungsdaten

Die Quantifizierung der Nachnutzung ist eine zentrale Herausforderung für die Einrichtungen und deren dauerhaften Fortbestand, denn: „Keine der Infrastrukturen für eine digitale Langzeitarchivierung lässt sich dauerhaft betreiben, wenn es keine Nutzer gibt.“ (Klump 2010: 111) Die Abschätzung, inwiefern sich eine Archivierung lohnt, wird jedoch dadurch, dass Teile des Materials mitunter erst zu einem deutlich späteren Zeitpunkt an

Relevanz, für eine Nachnutzung, gewinnen können (vgl. dgv 2018: 4) sowie, dass sich eine Kultur der Nachnutzung in vielen Disziplinen erst entwickeln muss, verkompliziert, sodass Nutzungszahlen mit Vorsicht betrachtet werden müssen.

Opitz und Mauers Befragung im Rahmen ihrer Machbarkeitsstudie zu Sekundärnutzung qualitativer sozialwissenschaftlicher Daten hatte zum Ergebnis, dass die befragten Wissenschaftler_innen mit Nachnutzungserfahrung angaben, dass über 40 Prozent die Daten für erneute Forschungsprojekte verwenden, ein Drittel sie in der Lehre einsetzen und ein Viertel jene für Qualifikationsarbeiten nutzen (vgl. 2005: 10).

Eine, zwar begrenzte, Möglichkeit die Nachnutzung und deren Auswirkungen zu messen, stellt die Messung von Downloadzahlen bzw. Zitationszahlen der Daten in anderen Publikationen durch Onlinerecherche sowie die Aufrufe der Seite, wo die Datensätze hinterlegt sind, dar, ebenso wie explorative Messmethoden wie die Erwähnung in sozialen Medien (vgl. Bishop/Kuula-Luumi 2017: 3, 7).

Wie bei quantitativen Daten gelten auch bei qualitativen Daten die von Ball und Duke beschriebenen Einschränkungen bei einem solchen Vorgehen, so sind manche Messungen sehr unpräzise, Doppelungen werden nicht immer erkannt und es besteht die Möglichkeit der Manipulation. Welche Methoden adäquat Auskunft geben, ist noch umstritten (vgl. Ball/Duke 2015: 12).

Einen Überblick können die Daten von Bishop und Kuula-Luumi bezüglich des langetablierten UK Data Service liefern: Dort wurde die Mehrzahl der Datensammlungen mindestens einmal verwendet. Die Sorge, dass große Datenmengen archiviert, aber nie angetastet werden, ist nachweislich unbegründet (vgl. 2017: 9). Ein kleiner Anteil der Sammlungen wurde am stärksten genutzt; ähnliche Nutzungsmuster wurden für die Verwendung quantitativer Daten beim UK Data Service festgestellt. Insgesamt wurden 42 % der Downloads von Postgraduierten, 25 % von Studierenden und 27 % von Mitarbeiter_innen von Hochschulen durchgeführt. Insgesamt wurden 15 % der Downloads für die Forschung, 13% für die Lehre und die deutliche Mehrheit, 64% für das Lernen verwendet (ebd.: 4).

Für Datengebende und Forschende ist es zudem interessant über die Art der Nutzung einen Überblick zu bekommen. „Für rund ein Viertel der Befragten stellt die Kontrolle über die weitere Nutzung eine wichtige Bedingung für die Bereitschaft zur Abgabe von Daten an ein

Archiv dar. Dazu gehört u. a. die Information, wer die Daten wofür nutzt oder aber auch die Genehmigung durch den Datengeber für jede einzelne Weitergabe.“ (Opitz/Mauer 2005: 16) Einen Lösungsansatz liefert UK Data Service mit dem Homepagebereich „data in use“, wo ausgewählte Anwendungen, die auf den im Archiv gesicherten Daten beruhen, dargestellt werden (<https://www.ukdataservice.ac.uk/use-data/data-in-use.aspx>).

In einer lange etablierten Einrichtung wachsen die Nutzungszahlen derzeit jährlich um 10 Prozent, wobei es sich bei der Sekundärnutzung in diesem Feld um eine anerkannte Methode handelt, insbesondere, weil eigene Daten zu erheben oft sehr ressourcenintensiv ist (vgl. TN7: 354-79). Sie problematisierte darüber hinaus, dass sich auf „ungefähr 20 Prozent dieser Daten [...] ca. 90 Prozent der Nutzung“ (ebd.: 511-12) konzentriert. Auch bei anderen Repositorien würde ein Großteil der Daten nicht genutzt werden. Downloadzahlen sollten nicht mit Nutzungen verwechselt werden. Tatsächliche Nutzung finde, nach einer sehr engen Definition, erst dann statt, wenn die Daten für die wissenschaftliche Beantwortung einer Fragestellung genutzt werden und publizierte Ergebnisse hervorbringt. „Das wird bei den allermeisten Daten [...] gar nicht der Fall sein [...] Und da stellt sich mir dann die Frage: Naja, wie viel Geld sollte man denn da rein investieren?“ (ebd.: 517-19)

Hier zeigt sich, dass sich die Einschätzungen, was als Nutzung gilt, zwischen den Einrichtungen unterscheiden, so fassen andere Institutionen z. B. die Anwendung in der Lehre (s. o.) ebenfalls darunter. Manche Einrichtungen können keine Angaben zur quantitativen Nutzung ihrer Bestände machen, da sie keinen Counter nutzen (vgl. TN4; TN1). Kritisch angemerkt wurde in einem Interview, dass dessen Einsatz problematisch sei, da dies dazu führen könne, dass bestimmte Studien nur angeklickt werden würden, weil sie bereits hohe Aufrufzahlen aufwiesen, was zu einer „self-fulfilling prophecy“ (TN1: 325) und der Schaffung „eine(r) Hitparade“ (ebd.: 328) führen könne. Problematisch sei die Quantifizierung auch deshalb, weil das Archiv noch über keinen sehr großen Datenbestand verfügt und dementsprechend nur für einen begrenzten Themenumfang Daten bereithält (ebd.: 409-13).

Die sehr selektive Nutzung des gesammelten Materials liege auch im Mangel an Nutzungsvorbildern begründet. Nutzende hätten rückgemeldet, dass das Fremdmaterial, anders als das Selbsterstellte, nicht mit dem gleichen Maß an Offenheit produziert worden sei, sodass es wenig Anchlüsse für alternative Fragestellungen bereithält (vgl. TN5). Eine Interviewte wies darauf hin, dass, je allgemeiner die Forschungen sind, desto mehr Nachnutzungspotenziale möglich sind. Aus diesem Grund wird bei der eigenen

Datenakquise versucht, möglichst breite Mehrthemenenerhebungen zu sichern (vgl. TN6: 130-41).

Bishop/Kuula-Luumi beschreiben, dass das größte Wiederverwendungspotenzial existiert, wenn das Thema, das für die Wiederverwendung von Interesse ist, in der ursprünglichen Forschung nicht im Mittelpunkt stand. Dies gilt insbesondere dann, wenn es schwierig sein kann, direkte Antworten zu erhalten, weil das Thema sensibel (Gesundheit), moralisch beladen (Lebensmittel) oder als selbstverständlich angesehen wird (Familienpraktiken) (vgl. 2017: 12).

4.5.1) Fachwissenschaftlicher Hintergrund und Ausbildung der Datenkurator_innen

Vielfach wurde in den Interviews zugleich betont, wie wichtig es ist, dass die Kuratierenden einen fachwissenschaftlichen Hintergrund besitzen (vgl. TN1: 670-671). Diese im weitesten Sinne sozialwissenschaftliche Ausbildung ist insbesondere für die Beratung von Nutzenden und zur (inhaltlichen) Prüfung, Aufbereitung und Dokumentation der Daten notwendig (vgl. TN7: 86-100).

Ein Archiv beschrieb, dass sie unterschiedliche Arbeitsabläufe mit verschiedenen ausgebildeten Mitarbeiter_innen, von Praktikant_innen über Bibliotheksmitarbeiter_innen bis hin zu Forschenden, ausprobiert hätten. Bei einfachen Prozessen gab es kaum Probleme bei der Datenaufbereitung. Je mehr Kontext und fachspezifisches Vorwissen z. B. bezüglich ethnografischer Bezeichnungen benötigt wurde, desto mehr Fehler entstanden, beispielsweise die Verwechslung von ethnischen Bezeichnungen mit geografischen. Beste Ergebnisse konnten mit fachkundigen Datenkurator_innen, die über wissenschaftliche Kenntnisse verfügten, erreicht werden. Die Forschenden selbst seien „ein bisschen zu betriebsblind oder haben auch nicht unbedingt die Geduld, um sehr genau an den Beständen zu arbeiten.“ (TN4: 138-139)

Trotz des breit aufgestellten fachlichen Hintergrunds der Angestellten konstatierte eine Einrichtung, dass eine enge Kooperation mit den Forschenden weiterhin nötig sei, da sie zurzeit nicht die Ressourcen besäßen, um über Expert_innen für jegliche Methode zu verfügen. Der teaminterne Austausch ist deshalb wichtig, da dort am konkreten Material kollektiv spezifische Lösungsansätze entwickelt werden können. Die Notwendigkeit fachlicher Expertise, insbesondere in der Beratung, „wird total unterschätzt“ (TN1: 691).

Eine beiläufige Behandlung von qualitativen Methoden sollte ebenso bei der Implementierung des Themas in Studiengänge vermieden werden. „Müssen wir [...] ein bisschen kämpfen, weil meistens quantitative Forschung höher gesetzt wird“ (ebd.: 694) und „wirklich mal auf Augenhöhe eigene Qualifikationen dafür“ (ebd.: 678) schaffen.

Auch in einem anderen Interview wurde darauf hingewiesen, dass das Forschungsdatenmanagement bei qualitativen Daten in der Lehre „nicht wirklich verankert“ (TN2: 261) ist. Angebote würden oft universitätsübergreifend unterbreitet und dabei qualitative Methoden und damit einhergehende Probleme zu wenig thematisiert werden (ebd.: 263-264). Um diese sichtbar zu machen und das Missverhältnis zu ändern, veranstaltet eine Einrichtung eigene Workshops, geht in die Lehre und kooperiert mit Hochschulen und Forschungsverbänden, „um [...] das Verständnis vielleicht (zu) erweitern und andere Perspektiven (zu)zulassen.“ (ebd.: 649-50) Die Wissensweitergabe an Forschende ist ebenfalls nötig, um so Hemmnisse und Vorurteile abzubauen, sodass sich Prozesse des Datenmanagements und der Sekundärnutzung normalisieren können. Obgleich viele Einrichtungen bereits praktische Arbeitsabläufe und -leitfäden für die Forschenden erstellt haben (vgl. TN3: 135-136), ist eine kompetente und kontinuierliche beratende Begleitung der Forschenden, insbesondere in Bereichen, die mit einem besonders hohen Maß an Unsicherheit einhergehen, wie rechtliche und forschungsethische Fragen, notwendig.

Problematisiert wurde zugleich der Mangel an Qualifikationsmöglichkeiten für Personen, die in der qualitativen sozialwissenschaftlichen Datenkuration arbeiten wollen. Die Einrichtungen: „können auch immer nur so effektiv und so gut arbeiten, wie die Leute qualifiziert sind“ (TN1: 661-662). Die Schaffung von Studiengängen und Weiterbildungsmöglichkeiten ist nötig, was bereits stattfindet. So beteiligt sich eine Einrichtung am „GO TRAIN“-Programm, wodurch Bewusstsein für die FAIR-Prinzipien und dahin gehende Qualifikationsmöglichkeiten geschaffen werden soll und die Einrichtung zukünftig dabei auf die Spezifik qualitativer sozialwissenschaftlicher Daten hinweisen will (vgl. TN1: 702-710).

Ziel ist: „to create a scalable, stakeholder-driven framework to rapidly train large numbers of competent, certified data stewards. A major challenge is to effectively coordinate the numerous and diverse stakeholders in FAIR-related training where currently fragmentation and re-invention of the wheel are potential risks. GO TRAIN seeks an open,

canonical training and education approach that can serve as a touchstone for global FAIR skills development drawing on agreed best practices.” (GFISCO 2020)²⁴

In einer Unterhaltung im Zuge des Erasmus+-Aufenthalts merkte eine Gesprächspartnerin die zu geringe Zahl von Universitätsveranstaltungen zum Thema Bilddaten an. Fachwissenschaftler_innen, die sich für eine Arbeit in diesem Bereich interessieren, müssten umfangreich eingearbeitet werden und für viele Probleme dann im nächsten Schritt individuelle Lösungen entwickelt werden (vgl. ebd.: 38-41).

4.5.2) Notwendigkeit der dauerhaften Finanzierung

Als zentrales Hindernis, für die Etablierung von Forschungsdatenmanagementinfrastrukturen, ist die Frage der dauerhaften Finanzierung (vgl. auch Hollstein/Strübing 2018c: 105). Insbesondere, weil der Aufbereitungsaufwand (Klärung von Rechtsfragen, Kontextualisierung etc.) bei vielen qualitativen sozialwissenschaftlichen Daten sehr umfangreich ist und sich das Nachnutzungspotenzial vielfach erst zu einem späteren Zeitpunkt einstellt, ist eine möglichst unbegrenzte Archivierung sinnvoll (vgl. Imeri et al. 2019: 33). Die Tätigkeiten der Einrichtungen sind Daueraufgaben und brauchen unbefristete Anstellungsverhältnisse und eine ausreichende Ressourcenausstattung, um diese Aufgaben sicherzustellen (vgl. TN1: 586-589). Der Archivierungsausbau auf Basis eines längeren finanziellen Planungshorizonts gilt als sinnvoll, da sonst nur wieder „Insellösungen“ (TN4: 13) geschaffen werden, „die kein Mensch verwenden kann, wie man es schon häufiger hatte.“ (ebd.) Eine ausreichende Personalausstattung wäre daneben auch nötig, um in der Lehre tätig werden zu können, regelmäßige Schulungen und eine intensive forschungsbegleitende Unterstützung, beispielsweise in Form von Datenmanagementplänen für Forschende, anbieten zu können (TN4: 299-306).

In einem Interview wurde davor gewarnt, dass bei einem Auslaufen der Projektfinanzierung bereits archivierte Daten verloren gehen könnten. So wurde die Frage aufgeworfen, ob es mit der NFDI zukünftig zentrale Repositorien gäbe: „Also den einen Punkt, wo ich gewährleisten kann, egal wie sich die Landschaft ändert, es gibt auf jeden Fall

²⁴ Vgl. hierzu auch die „Empfehlungen zu Berufs- und Ausbildungsperspektiven für den Arbeitsmarkt Wissenschaft“ des RfII, in welchen dieses Problemfeld angesprochen wird und Handlungsempfehlungen formuliert werden (vgl. RfII 2019: 1).

ein back up irgendwo.“ (TN6: 527-28) Zwar sei eine Verstetigung absehbar, da „es [...] wissenschaftspolitisch gewollt“ (TN1: 574) sei. „Aber in der Regel ist das ein langer Weg dahin.“ (ebd.: 588-89) Zugleich wurde eine Phase der strukturierten Bereinigung prognostiziert, da gerade „eine große Heterogenität, eine Parallelität, ein Nebeneinander- und Aneinander-Vorbei-Entwickeln (vorherrscht). Es ist [...] eine gewisse Goldgräberstimmung.“ (TN7: 447-48) „Aber es wird eine Abschwächung geben und eine Normalisierung. Forschungsdaten sind nicht alles.“ (ebd.: 456-57)

4.5.3) Materialauswahl und -auffindbarkeit

Neben der dauerhaften Finanzierung wurde zudem der Komplex der Materialauswahl und Datenauffindbarkeit als zentrale Herausforderung für die Zukunft des Forschungsdatenmanagements genannt. Insbesondere länger etablierte Einrichtungen, die bereits über einen größeren Datenbestand verfügen, erwähnten dies (vgl. TN1; TN7). Nutzer_innen stünden bei wachsendem Datenangebot zusehends vor der Schwierigkeit, einen umfassenden Überblick über den Bestand zu gewinnen. Um diese Auffindbarkeit sicherzustellen, ist zudem in manchen Forschungsbereichen die Etablierung von kontrolliertem Vokabular notwendig (vgl. TN3: 114).

In einem Interview wurde der Wunsch von Forschenden nach einer Verminderung von technischen Zugangshürden, beispielsweise via Virtualisierung oder in Form eines Remote-Access, angeführt. Da jedoch ein Großteil des Materials nicht ohne Beschränkungen öffentlich zugänglich gemacht werden kann, steigt bei erhöhter Zugangsgewährung auch die Notwendigkeit, Kontrollen und Autorisierungsmaßnahmen durchzuführen, die sich nicht automatisieren lassen (vgl. TN6: 512-13). Dafür müssen Lösungen gefunden und Ressourcen miteingeplant werden. Offenheit der Daten bedeutet auch hier grundsätzliche Zugänglichkeit, jedoch nur unter kontrollierten Bedingungen (vgl. Imeri 2019: 52).

4.5.4) Zwischenfazit hinsichtlich der Herausforderungen und Hindernisse in Bezug auf das Management qualitativer sozialwissenschaftlicher Forschungsdaten

Die Archivierung und Nachnutzung von qualitativen Daten im deutschsprachigen Raum befindet sich noch in der Etablierungsphase. Einige Archive verfügen noch nicht über einen umfangreichen Datenbestand, sodass sie nur für einen limitierten Themenbereich Daten zur

Verfügung stellen können. Nutzungszahlen und damit einhergehende Quantifizierungsverfahren sind aus diesem Grund nicht sehr aussagekräftig und müssen zum aktuellen Etablierungsstatus in Relation gesetzt werden. Mehrere Einrichtungen verwenden keinen Counter, sodass sie Downloadzahlen nicht bestimmen können (vgl. TN1; TN4). Erschwerend kommt hinzu, dass die befragten Einrichtungen unterschiedlich definieren, wann von einer tatsächlichen Nachnutzung gesprochen werden kann. Der reine Download ist nicht automatisch gleichbedeutend mit einer tatsächlichen Nutzung. Für manche Einrichtungen ist die Anwendung in der Lehre bereits eine Nutzung, für andere ist diese erst erfolgt, wenn die Nachnutzung die wissenschaftliche Publikation von Ergebnissen zur Folge hat.

Innerhalb der Interviews wurde hervorgehoben, dass die Datenkurator_innen über einen fachwissenschaftlichen Hintergrund verfügen und ausreichend ausgebildet sein sollten. Dies ist sowohl bei der Nutzendenberatung als auch bei der Datenprüfung, -aufbereitung und -dokumentation wichtig (vgl. TN1; TN4; TN7). Der Umgang mit qualitativen Daten sollte zudem stärker in der Lehre, die das Forschungsdatenmanagement umfasst, integriert werden und nicht nur am Rande thematisiert werden (vgl. TN2: 261). Die Schaffung von Qualifizierungsangeboten z. B. in Form von Studiengängen und berufsbegleitenden Weiterbildungsmöglichkeiten muss weiter ausgebaut werden.

Um eine dauerhafte und verlustsichere Archivierung der Daten sicherzustellen ist darüber hinaus eine kontinuierliche Finanzierung der Projekte, in Form von unbefristeten Stellenanteilen und ausreichend Sachmitteln notwendig. Ist dies nicht gewährleistet besteht die Gefahr, dass erneut nur „Insellösungen“ (TN4: 13) geschaffen werden. Personal bedarf es auch, um Lehrveranstaltungen, Schulungen und die forschungsbegleitende Unterstützung Forschender realisieren zu können.

Im Zuge wachsender Datenbestände wird die Frage danach, wie das Material zukünftig ausgewählt und auffindbar gemacht wird zusehends wichtiger. Es müssen Möglichkeiten etabliert werden die Daten überschaubar zu machen, was u. a. die Schaffung bzw. Anwendung von kontrolliertem Vokabular notwendig macht (vgl. TN3: 114).

5.) Fazit und Ausblick

Zusammenfassend ist auf Basis der Studie festzuhalten, dass sich aus der Archivierung, Bereitstellung und Nachnutzung qualitativer sozialwissenschaftlicher Forschungsdaten zahlreiche Chancen für die Forschung ergeben. Auf Ebene der Archivierung sind die Daten in professionellen Forschungsdatenzentren und -archiven am ehesten vor Verlust und missbräuchlicher Verwendung geschützt. Sie besitzen das Potenzial für vielfältige Nachnutzungsszenarien. Manche Einrichtungen verfügen über langjährige dahin gehende Erfahrungen, anderen Projekten fehlen derzeit noch die Mittel oder sie sind noch in der explorativen Erprobungsphase.

Die Materialien können als Datengrundlage und Inspirationsquelle für neue Forschungen sowie in der Methodenausbildung und Lehre genutzt werden. Hierbei geht es weniger um eine Kontrolle oder Reproduktion von Ergebnissen, sondern eher um das dauerhafte Bereitstellen für neue Analysen. Existierende Daten können effizienter genutzt und die kostenintensive wiederholte Durchführung thematisch ähnlicher Erhebungen reduziert werden. Die Erstellung von Datenmanagementplänen ermuntert die Primärforschenden sich intensiver mit Urheberrechts-, Datenschutz- und Persönlichkeitsschutzfragen zu befassen. Findet dieser Prozess forschungsbegleitend statt, kann gleichzeitig die Datenqualität erhöht werden. Managementpläne müssen jedoch aufgrund des vielfach dynamischen Forschungsprozesses ein hohes Maß an Flexibilität aufweisen. Es müssen noch umfassende Erfahrungswerte, inwiefern sich alle Materialien gleichermaßen nutzen lassen, und problemspezifische Lösungsansätze geschaffen werden (vgl. von Unger 2018: 96). Mitunter ist erst nach Abschluss der Forschung zu entscheiden, inwiefern das Material archiviert werden kann, um so insbesondere Formen der Selbstzensur der Forschenden in Antizipation einer Nachnutzung ihrer Daten, zu verhindern (vgl. Hollstein/Strübing 2018c: 103 f.).

Auf mehreren Ebenen besteht ein Spannungsfeld zwischen Standardisierungsversuchen einerseits, sehr diversen Forschungsdaten und -methoden sowie einem prinzipiell offenen Forschungsprozess andererseits (vgl. Imeri 2018b: 222). Ein Balanceakt zwischen individualisierten Lösungen und ressourcenarmen generischen Standardisierungsmöglichkeiten und Arbeitsroutinen bleibt deshalb nötig. Aufgrund der Spezifik des Materials und des notwendigen Fachwissens der Kuratierenden, ist die Archivierung in einem generischen Datenrepositorium oft nicht möglich. Zumal die

Einrichtungen einen kontrollierten und vielfach gestaffelten Zugang und adäquate Nutzungsverträge bereitstellen müssen.

Um eine Nachnutzung möglich zu machen, bedarf es oftmals ausführlicher Kontextualisierung via Datendokumentation. Auch hier gilt es zwischen Anforderungen des Daten- und Persönlichkeitsschutzes, der durch Anonymisierungen hergestellt werden soll und der Erhaltung ausreichender Interpretationsmöglichkeiten für Nachnutzungen zu vermitteln (vgl. Sterzer/Kretzer 2019: 114). Das Ideal der offenen Daten muss mit dem notwendigen Umfang an Quellenschutz austariert werden. Der Schutz der Forschenden, Beforschten und deren Vertrauensbeziehung hat Vorrang vor jeglicher Nachnutzung und die Weiterverarbeitung muss immer fallspezifisch entschieden werden, sodass „Archivierung [...] nicht automatisch auch Zugänglichkeit für Sekundärnutzungen“ (Hollstein/Strübing 2018b: 10) bedeutet. Die Akzeptanz bezüglich der Datenarchivierung hängt davon ab, ob dies gewährleistet wird und Forschungszugänge dadurch nicht erschwert werden. Insbesondere, weil Konsequenzen nicht immer gänzlich abgeschätzt und sich im Zeitverlauf ändern können, muss es einen flexiblen Umgang mit der Datenzugänglichkeit oder gar Löschung geben. Eine gestaffelte und kontrollierte Zugänglichkeit und rechtliche Vereinbarungen, die verhindern, dass nicht gewünschte Datennutzungen stattfinden und die Daten sicher sind, wenden die meisten Einrichtungen bereits an. Insbesondere bei kooperativer Datenproduktion müssen zudem praktikable Wege gefunden werden, um einerseits urheberrechtliche Arrangements zu finden, die die Interessen der Beforschten anerkennen und gleichzeitig Rechtssicherheit und Zugänglichkeit schaffen. Nicht nur rechtliche, sondern ebenso ethische Maßgaben sollten bei den Entscheidungen eine Rolle spielen. Sowohl bei rechtlich-ethischen Fragen als auch hinsichtlich der Kontextualisierung oder anderen Unsicherheiten werden Forschende durch die Einrichtungen aktiv beratend unterstützt.

Es muss sichergestellt werden, dass die Daten auffindbar sind, wozu die Entwicklung kontrollierter Vokabulare notwendig ist. Für (älteres) Material braucht es technische Verbesserungen zur einfacheren Suche auf der Ebene des einzelnen Datums.

Ob sich die Archivierungspotenziale und eine Kultur der Nachnutzung ähnlich wie bei quantitativen auch bei qualitativen sozialwissenschaftlichen Daten umfassend etablieren können, hängt von mehreren Faktoren ab: Einerseits, ob Sekundärforschungen als Bereicherung akzeptiert werden und inwiefern sich die Arbeit mit Sekundärmaterial in die

eigene Arbeitspraxis integrieren lässt. Und, ob Archivierungsleistungen als anerkannte eigenständige wissenschaftliche Leistung Würdigung erfahren, z. B. durch Zitationen anderer Forscher (vgl. Hollstein/Strübing 2018c: 104 f.). Dies ist auch davon abhängig, ob forschungsökonomische Machbarkeitseinwände, bezüglich des damit einhergehenden Arbeits- und Finanzierungsaufwands der Datenaufbereitung aus dem Weg geräumt werden können. Schließlich hängt die Etablierung daran, inwiefern all dies Teil der universitären Ausbildung wird.

Zentral ist eine Fachcommunityanbindung, um einerseits Anerkennung zu schaffen, Bedarfe zu kennen und zum anderen, um ein zukünftiges Regulativ zu schaffen, das gemeinsam mit den Dateninfrastrukturen Entscheidungskriterien dahingehend entwickeln kann, was archiviert werden soll. Es ist notwendig, dass sich die Einrichtungen mit den Fachgesellschaften und -communities weiterhin gemeinsam strategisch für eigene distinkte Lösungen einsetzen und mithelfen Standards zu setzen, da sonst die Gefahr droht, dass ihre fachspezifischen Bedarfe nicht berücksichtigt werden und Prozesse ohne ihre Mitwirkung ablaufen.

Zukünftig muss sich zeigen, welches Material sich gut für potenzielle Nachnutzungen eignet und inwiefern Forschende damit arbeiten werden. Dies ist jedoch nicht für alle erhobenen Daten sinnvoll und möglich. Der Methoden- und Materialreichtum bedeutet, dass vielfach unterschiedliche Lösungsansätze gefunden werden müssen, um Routinen zu entwickeln, wobei dies für qualitative Daten immer mit manuellem Arbeitsaufwand einhergehen wird. Es gibt kein Universalmodell, vielmehr müssen auf Basis der spezifischen Forschungsmethoden und -implikationen flexibel standardisierbare und modulare Lösungen erarbeitet werden. Dieser Prozess läuft bereits häufig vielversprechend ab, wobei für eine zukünftige Verstetigung sowohl eine fachliche Ausbildung von Datenkurator_innen, eine dauerhafte Finanzierung sowie Auswahlmechanismen dahingehend, was archiviert werden soll, nötig sind.

6.) Abkürzungsverzeichnis

ACUMEN	Academic careers understood through measurement and norms
DDC	Digital Curation Centre
DDI	Data Documentation Initiative
DFG	Deutsche Forschungsgemeinschaft
DGfE	Deutsche Gesellschaft für Erziehungswissenschaft
DGPs	Deutsche Gesellschaft für Psychologie
DGS	Deutsche Gesellschaft für Soziologie
DGSKA	Deutsche Gesellschaft für Sozial- und Kulturanthropologie
dgv	Deutsche Gesellschaft für Volkskunde
DSGVO	Datenschutz-Grundverordnung
EASA	European Association of Social Anthropologists
EK bzw. EC	Europäische Kommission bzw. European Commission
ELSST	European Language Social Science Thesaurus
EU	Europäische Union
FID	Fachinformationsdienst
Fachkollegium 106	Fachkollegium 106 Sozial- und Kulturanthropologie, Außereuropäische Kulturen, Judaistik und Religionswissenschaft
GFISCO	GO FAIR International Support & Coordination Office
GND	Gemeinsame Normdatei
HRK	Hochschulrektorenkonferenz
ISCO	International Standard Classification of Occupations
Kompetenznetz AHF	Kompetenznetz Angeborene Herzfehler
NFDI	Nationale Forschungsdateninfrastruktur
OECD	Organisation for Economic Co-operation and Development bzw. Organisation für wirtschaftliche Zusammenarbeit und Entwicklung
ÖAW ISA	Österreichische Akademie der Wissenschaften/Institut für Sozialanthropologie
RatSWD	Rat für Sozial- und Wirtschaftsdaten
RfII	Rat für Informationsinfrastrukturen
RDAIIDSIG	Research Data Alliance International Indigenous Data Sovereignty Interest Group
WR	Wissenschaftsrat

7.) Literatur- und Quellenverzeichnis

7.1) Literaturverzeichnis

Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen (2003). Online verfügbar unter https://openaccess.mpg.de/68053/Berliner_Erklaerung_dt_Version_07-2006.pdf, zuletzt geprüft am 7.4.2021.

ACUMEN Consortium (2014): Guidelines for Good Evaluation Practice with the ACUMEN Portfolio. Online verfügbar unter <http://www.research-acumen.eu/wp-content/uploads/Guidelines-Portfolio-TOC-document.v13dx.pdf>, zuletzt geprüft am 7.4.2021.

Allianz der deutschen Wissenschaftsorganisationen (2010): Grundsätze zum Umgang mit Forschungsdaten. In: RatSWD Working Paper 156/2010. Online verfügbar unter https://www.ratswd.de/download/RatSWD_WP_2010/RatSWD_WP_156.pdf, zuletzt geprüft am 7.4.2021.

Arbeitsgruppe Open-Access-Strategie Berlin (2015): Open-Access-Strategie für Berlin. Online verfügbar unter <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32016R0679&from=DE>, zuletzt geprüft am 7.4.2021.

Ball, Alex; Duke, Monica (2015): How to Track the Impact of Research Data with Metrics (DCC How-to Guides). Edinburgh: Digital Curation Centre (Hg.) Online verfügbar unter <https://www.dcc.ac.uk/guidance/how-guides>, zuletzt geprüft am 7.4.2021.

Ballsun-Stanton, Brian (2012): Asking About Data. Exploring Different Realities of Data via the Social Data Flow Network Methodology. Online verfügbar unter <https://philpapers.org/archive/BALAAD-3.pdf>, zuletzt geprüft am 7.4.2021.

Bambey, Doris; Meyermann, Alexia; Porzelt, Maïke; Rittberger, Marc (2018): Bereitstellung und Nachnutzung qualitativer Daten in der Bildungsforschung. Das Forschungsdatenzentrum (FDZ) Bildung am DIPF. In: Betina Hollstein und Jörg Strübing

(Hg.): Archivierung und Zugang zu qualitativen Daten. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD Working Paper 267/2018), S. 59–68.

Bishop, Libby; Kuula-Luumi, Arja (2017): Revisiting Qualitative Data Reuse: A Decade On. In: SAGE Open 7 (1), S. 1-15. DOI: 10.1177/2158244016685136.

Bogner, Alexander; Littig, Beate; Menz, Wolfgang (Hg.) (2005): Das Experteninterview – Theorie, Methode, Anwendung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Bogner, Alexander; Menz, Wolfgang (2005): Das theoriegenerierende Experteninterview - Erkenntnisinteresse, Wissensformen, Interaktion. In: Alexander Bogner, Beate Littig und Wolfgang Menz (Hg.): Das Experteninterview – Theorie, Methode, Anwendung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 33– 70.

Bogner, Alexander; Littig, Beate; Menz, Wolfgang (2014): Interviews mit Experten. Eine praxisorientierte Einführung. Wiesbaden: Springer VS (Lehrbuch). Online verfügbar unter <https://link.springer.com/book/10.1007/978-3-531-19416-5>, zuletzt geprüft am 7.4.2021.

Bohnsack, Ralf; Marotzki, Winfried; Meuser, Michael (Hg.) (2012): Hauptbegriffe Qualitativer Sozialforschung. Opladen/Farmington Hills: Verlag Barbara Budrich.

Borgman, Christine L. (2010): Research Data: Who will share what, with whom, when, and why? In: RatSWD Working Paper 161/2010. Online verfügbar unter https://www.konsortswd.de/wp-content/uploads/RatSWD_WP_161.pdf, zuletzt geprüft am 7.4.2021.

Breidenstein, Georg; Hirschauer, Stefan; Kalthoff, Herbert; Nieswand, Boris (2013): Ethnografie. Die Praxis der Feldforschung. Konstanz: UVK Verl.-Ges (UTB Sozialwissenschaften, Kulturwissenschaften).

Büttner, Stephan; Hobohm, Hans-Christoph (Hg.) (2011): Handbuch Forschungsdatenmanagement. Bad Honnef: Bock + Herchen.

Castillo, Rosa Cordillera A. (2018): Subverting 'formalised' ethics through mainstreaming critical research ethics and a responsive review process. In: Peter Pels, Igor Boog, J. Henrike Florusbosch, Zane Kripe, Tessa Minter, Metje Postma et al. (Hg.): Data management in anthropology: the next phase in ethics governance? In: *Social Anthropology/Anthropologie Sociale* (26), S. 406–407.

Corti, Louise (2018): 20 years of archiving and sharing qualitative data in the UK. In: Betina Hollstein und Jörg Strübing (Hg.): *Archivierung und Zugang zu qualitativen Daten*. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD Working Paper 267/2018), S. 14–25.

Corti, Louise; Fielding, Nigel (2016): Opportunities From the Digital Revolution. In: *SAGE Open* 6 (4), S. 1-13. DOI: 10.1177/2158244016678912.

Corti, Louise; Fielding, Nigel; Bishop, Libby (2016): Editorial for Special Edition, Digital Representations: Re- Using and Publishing Digital Qualitative Data. In: *SAGE Open* 6 (4), S. 1-3. DOI: 10.1177/2158244016678911.

Degkwitz, Andreas (Hg.) (2016): *Bibliothek der Zukunft. Zukunft der Bibliothek*. Festschrift für Elmar Mittler. Berlin, Boston: De Gruyter.

Degkwitz, Andreas (2019): Vorwort. In: Sabine Imeri; Wjatscheslaw Sterzer; Matthias Harbeck: *Forschungsdatenmanagement in den ethnologischen Fächern*. In: *Schriftenreihe der Universitätsbibliothek der Humboldt-Universität zu Berlin* (67), S. 6-7.

Depping, Ralf (2020): SocioHub (Fachinformationsdienst Soziologie) als Teil der überregionalen Forschungsdateninfrastruktur. Online verfügbar unter https://www.ratswd.de/dl/media/Depping8KSWD_FID.pdf, zuletzt geprüft am 7.4.2021.

Deutsche Forschungsgemeinschaft (DFG) (2009): *Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten*. Online verfügbar unter https://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901.pdf, zuletzt geprüft am 7.4.2021.

Deutsche Forschungsgemeinschaft (DFG) (2013 [1998]): Sicherung guter wissenschaftlicher Praxis / Safeguarding Good Scientific Practice. Online verfügbar unter https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf, zuletzt geprüft am 7.4.2021.

Deutsche Forschungsgemeinschaft (DFG) (2015): Leitlinien zum Umgang mit Forschungsdaten. Online verfügbar unter http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf, zuletzt geprüft am 7.4.2021.

Deutsche Forschungsgemeinschaft (DFG) (2018): Leitfaden für die Antragstellung. Projektanträge. Online verfügbar unter https://www.dfg.de/formulare/54_01/54_01_de.pdf, zuletzt geprüft am 7.4.2021.

Deutsche Gesellschaft für Erziehungswissenschaft (DGfE) (2017): Stellungnahme der DGfE zur Archivierung, Bereitstellung und Nachnutzung qualitativer Forschungsdaten in der Erziehungswissenschaft. Online verfügbar unter http://www.dgfe.de/fileadmin/Ordner_Redakteure/Stellungnahmen/2017.09_Archivierung_qual._Daten.pdf, zuletzt geprüft am 7.4.2021.

Deutsche Gesellschaft für Sozial- und Kulturanthropologie (DGSKA) (2019): Positionspapier zum Umgang mit ethnologischen Forschungsdaten, Deutsche Gesellschaft für Sozial- und Kulturanthropologie (DGSKA). Online verfügbar unter https://www.dgska.de/wp-content/uploads/2019/11/Positionspapier_Bearbeitet-fu%CC%88r-MV_24.09.2019.pdf, zuletzt geprüft am 7.4.2021.

Deutsche Gesellschaft für Soziologie (DGS) (2019): Bereitstellung und Nachnutzung von Forschungsdaten in der Soziologie. Stellungnahme des Vorstands und Konzils der DGS. Online verfügbar unter https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/dgs_stellungnahme_forschungsdaten.pdf, zuletzt geprüft am 7.4.2021.

Deutsche Gesellschaft für Volkskunde (dgv) (2018): Positionspapier zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten. Online verfügbar unter

https://www.d-g-v.de/wp-content/uploads/2019/04/dgv-Positionspapier_FDM.pdf, zuletzt geprüft am 7.4.2021.

Digital Curation Centre (DDC) (2020): DCC Curation Lifecycle Model. Online verfügbar unter <https://www.dcc.ac.uk/sites/default/files/documents/publications/DCCLifecycle.pdf>, zuletzt geprüft am 7.4.2021.

Dilger, Hansjörg (2017): Ethics, Epistemology and Ethnography: The Need for an Anthropological Debate on Ethical Review Processes in Germany. In: *Sociologus* 67 (2), S. 191–208. DOI: 10.3790/soc.67.2.191.

Dunkel, Wolfgang; Hanekop, Heidemarie (2018): FDZ für qualitative Forschungsdaten in der Arbeits- und Industriesoziologie: das interdisziplinäre Zentrum eLabour. In: Betina Hollstein und Jörg Strübing (Hg.): *Archivierung und Zugang zu qualitativen Daten*. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD Working Paper 267/2018), S. 36–46.

Eberhard, Igor; Kraus, Wolfgang (2018): Der Elefant im Raum. Ethnographisches Forschungsdatenmanagement als Herausforderung für Repositorien. In: *Mitteilungen der VÖB* 71 (1), S. 41–52. DOI: 10.31263/voebm.v71i1.2018.

Europäische Kommission (EK) (2012): Pressemitteilung: Wissenschaftliche Daten: freier Zugang zu Forschungsergebnissen wird Innovationskapazität der EU stärken. Online verfügbar unter https://ec.europa.eu/commission/presscorner/detail/de/IP_12_790, zuletzt geprüft am 7.4.2021.

Europäische Union (EU) (2016): Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung). Online verfügbar unter <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32016R0679%20&from=DE>, zuletzt geprüft am 7.4.2021.

European Association of Social Anthropologists (EASA) (2018): EASA's Statement on Data Governance in Ethnographic Projects. Online verfügbar unter

<https://www.easaonline.org/downloads/support/EASA%20statement%20on%20data%20governance.pdf>, zuletzt geprüft am 7.4.2021.

European Commission (EC) (2017): H2020 Programme. Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. Version 3.2. Online verfügbar unter https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf, zuletzt geprüft am 7.4.2021.

Fachkollegium 106 (2019): Handreichung des Fachkollegiums 106 Sozial- und Kulturanthropologie, Außereuropäische Kulturen, Judaistik und Religionswissenschaft zum Umgang mit Forschungsdaten. Online verfügbar unter https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/handreichung_fachkollegium_106_forschungsdaten.pdf, zuletzt geprüft am 7.4.2021.

Fecher, Benedikt; Friesike, Sascha (2013): Open Science: One Term, Five Schools of Thought. In: SSRN Journal. DOI: 10.2139/ssrn.2272036.

Finetti, Marco; Himmelrath, Armin (1999): Der Sündenfall. Betrug und Fälschung in der deutschen Wissenschaft. Stuttgart: Raabe (DUZ-Edition).

Flick, Uwe; Kardorff, Ernst von; Steinke, Ines (Hg.) (2005): Qualitative Sozialforschung. Ein Handbuch. Hamburg: Rowohlt Taschenbuch Verlag.

Flick, Uwe; Kardorff, Ernst von; Steinke, Ines (Hg.) (2012): Qualitative Forschung - Ein Handbuch. Hamburg: Rowohlt Taschenbuch Verlag.

Garz, Detlef; Kraimer, Klaus (Hg.) (1991): Qualitativ-empirische Sozialforschung: Konzepte, Methoden, Analysen. Opladen: Westdeutscher Verlag.

Gebel, Tobias; Grenzer, Matthis; Kreusch, Julia; Liebig, Stefan; Schuster, Heidi; Tschewinka, Ralf et al. (2015): Verboten ist, was nicht ausdrücklich erlaubt ist: Datenschutz in qualitativen Interviews. In: Forum Qualitative Sozialforschung 16 (2). Online verfügbar

unter <http://www.qualitative-research.net/index.php/fqs/%20article/viewFile/2266/3822>,
zuletzt geprüft am 7.4.2021.

Gläser, Jochen; Laudel, Grit (2010): Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen. Wiesbaden: VS Verlag.

Göttsch, Silke; Lehman, Albrecht (Hg.) (2007): Methoden der Volkskunde. Positionen, Quellen, Arbeitsweisen der Europäischen Ethnologie. Berlin: Dietrich Reimer Verlag.

Hahn, Hans Peter (2014): Ethnologie. Eine Einführung. Berlin: Suhrkamp.

Harbeck, Matthias; Imeri, Sabine; Sterzer, Wjatscheslaw (2018): Feldnotizen und Videomitschnitte. Zum Forschungsdatenmanagement qualitativer Daten am Beispiel der ethnologischen Fächer. In: o-bib. Das offene Bibliotheksjournal (Bd.5, Nr. 2), S. 123–141. DOI: 10.5282/O-BIB/2018H2S123-141.

Harbeck, Matthias; Kaun, Matthias (2019): Forschungsdaten und Fachinformationsdienste – eine Bestandsaufnahme. In: Preprints der Zeitschrift BIBLIOTHEK – Forschung und Praxis. DOI: 10.18452/19693.

Hirschauer, Stefan (2014): Sinn im Archiv? Zum Verhältnis von Nutzen, Kosten und Risiken der Datenarchivierung. In: Soziologie 43 (3), S. 300–312. Online verfügbar unter <https://publikationen.sozioologie.de/index.php/soziologie/article/view/795/1052>, zuletzt geprüft am 7.4.2021.

Hochschulrektorenkonferenz (HRK) (2015): Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können. Orientierungspfade, Handlungsoptionen, Szenarien. Empfehlung der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel. Online verfügbar unter https://www.hrk.de/fileadmin/redaktion/hrk/02-Dokumente/02-10-Publikationsdatenbank/Beitr-2016-01_Forschungsdatenmanagement.pdf, zuletzt geprüft am 7.4.2021.

Hollstein, Betina; Strübing, Jörg (Hg.) (2018a): Archivierung und Zugang zu qualitativen Daten. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD Working Paper 267/2018). DOI: 10.17620/02671.35.

Hollstein, Betina; Strübing, Jörg (2018b): Archivierung und Zugang zu qualitativen Daten. In: Betina Hollstein und Jörg Strübing (Hg.): Archivierung und Zugang zu qualitativen Daten. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD Working Paper 267/2018), S. 1–13.

Hollstein, Betina; Strübing, Jörg (2018c): Zentrale Ergebnisse und Empfehlungen aus dem Workshop für Archivierung und Sekundäranalyse qualitativer Forschungsdaten. In: Betina Hollstein und Jörg Strübing (Hg.): Archivierung und Zugang zu qualitativen Daten. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD Working Paper 267/2018), S. 101–105.

Hopf, Christel (2012): Qualitative Interviews - ein Überblick. In: Uwe Flick, Ernst von Kardorff und Ines Steinke (Hg.): Qualitative Forschung - Ein Handbuch. Hamburg: Rowohlt Taschenbuch Verlag, S. 349–360.

Huber, Elisabeth (2019a): Affektive Dimensionen von Forschungsdaten, ihrer Nachnutzung und Verwaltung. In: Working Paper SFB 1171 Affective Societies 01/19. Online verfügbar unter https://refubium.fu-berlin.de/bitstream/handle/fub188/24721/SFB1171_WP_01-19_Huber.pdf, zuletzt geprüft am 7.4.2021.

Huber, Elisabeth (2019b): Beratung und Entscheidungsfindung bei rechtlichen und ethischen Fragen des Forschungsdatenmanagements. In: Bausteine Forschungsdatenmanagement. Empfehlungen und Erfahrungsberichte für die Praxis von Forschungsdatenmanagerinnen und -managern (2), S. 31–39. DOI: 10.17192/BFDM.2019.2.8072.

Imeri, Sabine (2017): Open Data? Zum Umgang mit Forschungsdaten in den ethnologischen Fächern. In: Kratzke, Jonas; Heuveline, Vincent (Hg.) (2017): E-Science-Tage 2017 - Forschungsdaten managen, S. 167– 178.

Imeri, Sabine (2018a): Archivierung und Verantwortung. Zum Stand der Debatte über den Umgang mit Forschungsdaten in den ethnologischen Fächern. In: Betina Hollstein und Jörg Strübing (Hg.): Archivierung und Zugang zu qualitativen Daten. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD Working Paper 267/2018), S. 69–79.

Imeri, Sabine (2018b): Ordnen, archivieren, teilen. Forschungsdaten in den ethnologischen Fächern. In: Österreichische Zeitschrift für Volkskunde LXXII/121 (2), S. 213–243.

Imeri, Sabine (2019): ‚Open Data‘ in den ethnologischen Fächern - Möglichkeiten und Grenzen eines Konzepts. In: Klingner, Jens; Lühr, Merve (Hg.) (2019): Forschungsdesign 4.0 - Datengenerierung und Wissenstransfer in interdisziplinärer Perspektive, S. 45–59.

Imeri, Sabine; Danciu, Ida [Mitarb.] (2017): Open Data. Forschungsdatenmanagement in den ethnologischen Fächern. Auswertung einer Umfrage des Fachinformationsdienstes Sozial- und Kulturanthropologie an der Universitätsbibliothek der Humboldt-Universität zu Berlin 2016. Teil I: Statistiken. Online verfügbar unter http://www.evifa.de/cms/fileadmin/uploads/Umfrage_Bericht_Statistiken_1.0_14-06-2017.pdf, zuletzt geprüft am 7.4.2021.

Imeri, Sabine; Sterzer, Wjatscheslaw; Harbeck, Matthias (2019): Forschungsdatenmanagement in den ethnologischen Fächern. In: Schriftenreihe der Universitätsbibliothek der Humboldt-Universität zu Berlin (67). DOI: 10.18452/20090.

Jacobs, James A.; Humphrey, Charles (2004): Preserving research data. In: Commun. ACM 47 (9), S. 27. DOI: 10.1145/1015864.1015881.

Jensen, Uwe; Katsanidou Alexia; Zenk-Möltgen, Wolfgang (2011): Metadaten und Standards. In Büttner, Stephan und Hans-Christoph Hobohm (Hg.): Handbuch Forschungsdatenmanagement. Bad Honnef: Bock + Herchen, S. 83-100.

Klingner, Jens; Lühr, Merve (Hg.) (2019): Forschungsdesign 4.0 - Datengenerierung und Wissenstransfer in interdisziplinärer Perspektive. DOI: 10.25366/2019.04.

Kaden, Ben (2018): Warum Forschungsdaten nicht publiziert werden. In: LIBREAS. Library Ideas (33), S. 1–8. DOI: 10.18452/19284.

Kindling, Maxi; Burger, Marleen; Liebenau, Lisa; Lienhard, Claudia; Lilienthal, Svantje; Plewka, Paulina et al. (2013): Forschungsdatenmanagement an Hochschulen. Internationaler Überblick und Aspekte eines Konzepts für das Forschungsdatenmanagement an der Humboldt-Universität zu Berlin. DOI: 10.18452/13561.

Kindling, Maxi; Schirmbacher, Peter (2013): „Die digitale Forschungswelt“ als Gegenstand der Forschung. Lehrstuhl Informationsmanagement. In: Information, Wissenschaft & Praxis 64 (2-3), S. 127–136. DOI: 10.1515/iwp-2013-0017.

Klump, Jens (2010): Digitale Forschungsdaten. In: Heike Neuroth, Achim Oßwald, Regine Scheffel, Stefan Strathmann und Karsten Huth (Hg.): nestor-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, S. 104–115. Online verfügbar unter http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf, zuletzt geprüft am 7.4.2021.

Knoblauch, Herbert (2012): Transkription. In: Ralf Bohnsack, Winfried Marotzki und Michael Meuser (Hg.): Hauptbegriffe Qualitativer Sozialforschung. Opladen/Farmington Hills: Verlag Barbara Budrich, S. 159–160.

Knoblauch, Hubert; Wilke, René (2018): Forschungsdateninfrastrukturen für audio-visuelle Daten der Qualitativen Sozialforschung – Bedarf und Anforderungen. In: Betina Hollstein und Jörg Strübing (Hg.): Archivierung und Zugang zu qualitativen Daten. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD Working Paper 267/2018), S. 47–58.

Koenig, Robert (1997): Panel calls falsification in German case 'unprecedented'. In: Science (New York, N.Y.) 277 (5328), S. 894. DOI: 10.1126/science.277.5328.894.

Kowal, Sabine; O'Connel, Daniel C. (2005): Zur Transkription von Gesprächen. In: Uwe Flick, Ernst von Kardorff und Ines Steinke (Hg.): Qualitative Sozialforschung. Ein Handbuch. Hamburg: Rowohlt Taschenbuch Verlag, S. 437–447.

Kratzke, Jonas; Heuveline, Vincent (Hg.) (2017): E-Science-Tage 2017 - Forschungsdaten managen. DOI: 10.11588/HEIBOOKS.285.377.

Kuschel, Linda (2018): Wem "gehören" Forschungsdaten? In: *Forschung & Lehre* (9/2018), S. 764–766. Online verfügbar unter <https://www.forschung-und-lehre.de/wem-gehoren-forschungsdaten-1013/>, zuletzt geprüft am 7.4.2021.

Lederman, Rena (2018): Doing anthropology ethically takes practice: a US perspective on formalisation. In: Peter Pels, Igor Boog, J. Henrike Florusbosch, Zane Kripe, Tessa Minter, Metje Postma et al. (Hg.): *Data management in anthropology: the next phase in ethics governance?*, In: *Social Anthropology/Anthropologie Sociale* (26), 3, S. 408–409.

Leopold, Robert (2008): The second life of ethnographic fieldnotes. In: *ateliers* (32). DOI: 10.4000/ateliers.3132.

Medjedović, Irena; Witzel, Andreas (2010): *Wiederverwendung qualitativer Daten. Archivierung und Sekundärnutzung qualitativer Interviewtranskripte*. Wiesbaden: VS Verlag für Sozialwissenschaften (GWV).

Meier zu Verl, Christian; Meyer, Christian (2018): Probleme der Archivierung und sekundären Nutzung ethnografischer Daten. In: Betina Hollstein und Jörg Strübing (Hg.): *Archivierung und Zugang zu qualitativen Daten*. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD Working Paper 267/2018), S. 80–90.

Meuser, Michael; Nagel, Ulrike (1991): ExpertInneninterviews - vielfach erprobt, wenig bedacht: ein Beitrag zur qualitativen Methodendiskussion. In: Detlef Garz und Klaus Kraimer (Hg.): *Qualitativ-empirische Sozialforschung: Konzepte, Methoden, Analysen*. Opladen: Westdeutscher Verlag, S. 441–471.

Meuser, Michael; Nagel, Ulrike (2009): *Das Experteninterview – konzeptionelle Grundlagen und methodische Anlage*. In: Susanne Pickel, Detlef Jahn, Hans-Joachim Lauth und Gert Pickel (Hg.): *Methoden der vergleichenden Politik- und Sozialwissenschaft. Neue Entwicklungen und Anwendungen*. 1. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden, S. 465–480.

Neuroth, Heike; Oßwald, Achim; Scheffel, Regine; Strathmann, Stefan; Huth, Karsten (Hg.) (2010): nestor-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Online verfügbar unter http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf, zuletzt geprüft am 7.4.2021.

Opitz, Diane; Mauer, Reiner (2005): Erfahrungen mit der Sekundärnutzung von qualitativem Datenmaterial – Erste Ergebnisse einer schriftlichen Befragung im Rahmen der Machbarkeitsstudie zur Archivierung und Sekundärnutzung qualitativer Interviewdaten. In: Forum Qualitative Sozialforschung / Forum: Qualitative Social Research Vol. 6 (No. 1). DOI: 10.17169/FQS-6.1.510.

Organisation for economic co-operation and development (OECD) (2004): Science, Technology and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 - Final Communiqué. Online verfügbar unter <https://www.oecd.org/sti/sciencetechnologyandinnovationforthe21stcenturymeetingoftheoecdcommitteeforscientificandtechnologicalpolicyatministeriallevel29-30january2004-finalcommunique.htm>, zuletzt geprüft am 7.4.2021.

Organisation for economic co-operation and development (OECD) (2007): OECD Principles and Guidelines for Access to Research Data from Public Funding. Online verfügbar unter https://www.oecd-ilibrary.org/science-and-technology/oecd-principles-and-guidelines-for-access-to-research-data-from-public-funding_9789264034020-en-fr, zuletzt geprüft am 7.4.2021.

Österreichische Akademie der Wissenschaften/Institut für Sozialanthropologie (ÖAW ISA): Geschichte der Ethnographien Asiens. Online verfügbar unter <https://www.oeaw.ac.at/isa/forschung/translokale-themen/geschichte-der-ethnographien-asiens/>, zuletzt geprüft am 7.4.2021.

Pels, Peter (2018a): Data management in anthropology: the next phase in ethics governance? In: Peter Pels, Igor Boog, J. Henrike Florusbosch, Zane Kripe, Tessa Minter, Metje Postma et al. (Hg.): Data management in anthropology: the next phase in ethics governance?, In: Social Anthropology/Anthropologie Sociale (26), 3, S. 391–396.

Pels, Peter (2018b): Response (to Ethics of data management: an anthropological archaeology perspective by Heather Richards-Rissetto). In: Peter Pels, Igor Boog, J. Henrike Florusbosch, Zane Kripe, Tessa Minter, Metje Postma et al. (Hg.): Data management in anthropology: the next phase in ethics governance?, In: *Social Anthropology/Anthropologie Sociale* (26), 3, S. 411–412.

Pels, Peter; Boog, Igor; Henrike Florusbosch, J.; Kripe, Zane; Minter, Tessa; Postma, Metje et al. (Hg.) (2018a): Data management in anthropology: the next phase in ethics governance?, In: *Social Anthropology/Anthropologie Sociale* (26), 3, S. 391–413.

Pels, Peter; Boog, Igor; Florusbosch, J. Henrike; Kripe, Zane; Minter, Tessa; Postma, Metje (2018b): Data management for anthropologists and ethnographers: a position paper. In: Peter Pels, Igor Boog, J. Henrike Florusbosch, Zane Kripe, Tessa Minter, Metje Postma et al. (Hg.): Data management in anthropology: the next phase in ethics governance?, In: *Social Anthropology/Anthropologie Sociale* (26), 3, S. 397–402.

Pfadenhauer, Michaela (2005): Auf gleicher Augenhöhe reden – Das Experteninterview – ein Gespräch zwischen Experte und Quasi-Experte. In: Alexander Bogner, Beate Littig und Wolfgang Menz (Hg.): *Das Experteninterview – Theorie, Methode, Anwendung*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 113– 130.

Pickel, Susanne; Jahn, Detlef; Lauth, Hans-Joachim; Pickel, Gert (Hg.) (2009): *Methoden der vergleichenden Politik- und Sozialwissenschaft. Neue Entwicklungen und Anwendungen*. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden.

Queckbörner, Boris (2019): Forschungsdatenmanagement und Forschungsdaten in der Geschichtswissenschaft: Gegenwärtige Praxis und Perspektiven am Beispiel ausgewählter Sonderforschungsbereiche. In: *Berliner Handreichungen zur Bibliotheks- und Informationswissenschaft* 441. DOI: 10.18452/20460.

Rat für Informationsinfrastrukturen (RfII) (2019): *Digitale Kompetenzen - dringend gesucht! Empfehlungen zu Berufs- und Ausbildungsperspektiven für den Arbeitsmarkt*

Wissenschaft. Online verfügbar unter <http://www.rfii.de/?wpdmdl=3883>, zuletzt geprüft am 7.4.2021.

Rat für Sozial- und Wirtschaftsdaten (RatSWD) (2018a): Workshop „Archivierung und Zugang zu qualitativen Daten“. Online verfügbar unter https://www.ratswd.de/dl/media/Programm_RatSWD_Workshop_Qualitative_Daten_April_2018_Bremen.pdf, zuletzt geprüft am 7.4.2021.

Rat für Sozial- und Wirtschaftsdaten (RatSWD) (2018b): Forschungsdatenmanagement in den Sozial-, Verhaltens- und Wirtschaftswissenschaften. – Orientierungshilfen für die Beantragung und Begutachtung datengenerierender und datennutzender Forschungsprojekte. Berlin: RatSWD Output 3 (5). DOI: 10.17620/02671.7.

Research Data Alliance International Indigenous Data Sovereignty Interest Group (RDAIIDSIG) (2019): CARE Principles for Indigenous Data Governance. Online verfügbar unter https://static1.squarespace.com/static/5d3799de845604000199cd24/t/5da9f4479ecab221ce848fb2/157141%209335217/CARE+Principles_One+Pagers+FINAL_Oct_17_2019.pdf, zuletzt geprüft am 7.4.2021.

Rinsdorf, Lars (2013): Qualitative Methoden. In: Umlauf, Konrad; Seadle, Michael S.; Hauke, Petra; Fühles-Ubach, Simone (Hg.): Handbuch Methoden der Bibliotheks- und Informationswissenschaft. Bibliotheks-, Benutzerforschung, Informationsanalyse. Berlin, Boston: De Gruyter Saur, S. 64–79.

Rümpel, Stefanie (2011): Der Lebenszyklus von Forschungsdaten. In: Stephan Büttner und Hans-Christoph Hobohm (Hg.): Handbuch Forschungsdatenmanagement. Bad Honnef: Bock + Herchen, S. 25–34.

Sahle, Patrick; Kronenwett, Simone (2013): Jenseits der Daten. Überlegungen zu Datenzentren für die Geisteswissenschaften am Beispiel des Kölner ‘Data center for the humanities’. In: LIBREAS. Library Ideas (23). DOI: 10.18452/9043.

Schmidt, Christiane (2012): Analyse von Leitfadeninterviews. In: Uwe Flick, Ernst von Kardorff und Ines Steinke (Hg.): Qualitative Forschung - Ein Handbuch. Hamburg: Rowohlt Taschenbuch Verlag, S. 447–456.

Schmidt-Lauber, Brigitta (2007): Das qualitative Interview oder: Die Kunst des Reden-Lassens. In: Silke Göttisch und Albrecht Lehman (Hg.): Methoden der Volkskunde. Positionen, Quellen, Arbeitsweisen der Europäischen Ethnologie. Berlin: Dietrich Reimer Verlag, S. 169–188.

Scholze, Frank; Bertelmann, Roland; Kindling, Maxi; Pampel, Heinz; Vierkant, Paul (2016): Open Access und Forschungsdaten. In: Andreas Degkwitz (Hg.): Bibliothek der Zukunft. Zukunft der Bibliothek. Festschrift für Elmar Mittler. Berlin, Boston: De Gruyter, S. 156–164.

Schönbrodt, Felix; Gollwitzer, Mario; Abele-Brehm, Andrea (2017): Der Umgang mit Forschungsdaten im Fach Psychologie: Konkretisierung der DFG-Leitlinien. Im Auftrag des DGPs Vorstands. In: Psychologische Rundschau 68 (1), S. 20–35. DOI: 10.1026/0033-3042/a000341.

Six-Hohenbalken, Maria; Mehmet, Emir (2016): Abenteuer, Ethnograph, Pferdenarr. Online verfügbar unter <https://augustin.or.at/abenteurer-ethnograph-pferdenarr/>, zuletzt geprüft am 7.4.2021.

Sleeboom-Faulkner, Margaret; Simpson, Bob (2018): Research ethics and data management: perspectives from the UK. In: Peter Pels, Igor Boog, J. Henrike Florusbosch, Zane Kripe, Tessa Minter, Metje Postma et al. (Hg.): Data management in anthropology: the next phase in ethics governance?, In: Social Anthropology/Anthropologie Sociale (26), 3, S. 402–404.

Smioski, Andrea (2011): Wegweiser qualitative Datenarchivierung: Infrastruktur, Datenakquise, Dokumentation und Weitergabe. In: SWS-Rundschau 51 (2), S. 219–238. Online verfügbar unter http://www.sws-rundschau.at/archiv/SWS_2011_2_Smioski.pdf, zuletzt geprüft am 1.6.2020.

Smioski, Andrea (2013): Archivierungsstrategien für qualitative Daten. In: Forum Qualitative Sozialforschung 14 (3). DOI: 10.17169/FQS-14.3.1958.

Spindler, Gerald; Hillegeist, Tobias (2011): Rechtliche Probleme der elektronischen Langzeitarchivierung von Forschungsdaten. In: Stephan Büttner und Hans-Christoph Hobohm (Hg.): Handbuch Forschungsdatenmanagement. Bad Honnef: Bock + Herchen, S. 63–70.

Sterzer, Wjatscheslaw; Kretzer, Susanne (2019): Archivierungsstrategien anpassen – Herausforderungen und Lösungen für die Archivierung und Sekundärnutzung von ethnologischen Forschungsdaten. In: Bibliothek Forschung und Praxis 43 (1), S. 110–117. DOI: 10.1515/bfp-2019-2011.

The Royal Society (2012): Science as an open enterprise. The Royal Society Science Policy Centre report. Online verfügbar unter <https://royalsociety.org/-/media/policy/projects/sape/2012-06-20-saoe.pdf>, zuletzt geprüft am 7.4.2021.

Umlauf, Konrad; Seadle, Michael S.; Hauke, Petra; Fühles-Ubach, Simone (2013) (Hg.): Handbuch Methoden der Bibliotheks- und Informationswissenschaft. Bibliotheks-, Benutzerforschung, Informationsanalyse. Berlin, Boston: De Gruyter Saur.

Unger, Hella von (2018): Archivierung und Nachnutzung qualitativer Daten aus forschungsethischer Perspektive. In: Betina Hollstein und Jörg Strübing (Hg.): Archivierung und Zugang zu qualitativen Daten. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD Working Paper 267/2018), S. 91–100.

Voß, Jakob (2013): Was sind eigentlich Daten? In: LIBREAS Library Ideas (23). Online verfügbar unter <https://libreas.eu/ausgabe23/02voss/>, zuletzt geprüft am 7.4.2021.

Werner, Petra (2013): Qualitative Befragungen. In: Umlauf, Konrad; Seadle, Michael S.; Hauke, Petra; Fühles-Ubach, Simone (Hg.): Handbuch Methoden der Bibliotheks- und Informationswissenschaft. Bibliotheks-, Benutzerforschung, Informationsanalyse. Berlin, Boston: De Gruyter Saur, S. 128–151.

Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, I. Jsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie et al. (2016): The FAIR Guiding Principles for scientific data management and stewardship. In: Scientific data 3. DOI: 10.1038/sdata.2016.18.

Wissenschaftsrat (WR) (2012): Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020. Online verfügbar unter https://www.wissenschaftsrat.de/download/archiv/2359-12.pdf;jsessionid=2DD190BF2468DAB84F65E52%2038C5685A4.delivery1-master?__blob=publicationFile&v=3, zuletzt geprüft am 7.4.2021.

7.2) Quellenverzeichnis

GO FAIR International Support & Coordination Office (GFISCO): GO TRAIN. Online verfügbar unter <https://www.go-fair.org/go-fair-initiative/go-train/>, zuletzt geprüft am 7.4.2021.

Kompetenznetz Angeborene Herzfehler: Data Access Committee. Online verfügbar unter <https://www.kompetenznetz-ahf.de/wir/data-access-committee>, zuletzt geprüft am 7.4.2021.

UK Data Service: Data in use. "How our data are being used to advance research, inform policy and improve teaching". Online verfügbar unter <https://www.ukdataservice.ac.uk/use-data/data-in-use.aspx>, zuletzt geprüft am 7.4.2021.

8.) Anhang

8.1) Exemplarische Leitfadenübersicht

Die hier dargestellten Fragen wurden, je nach Kontext, spezifisch bezüglich Abfolge und Fokus an das zu interviewende Gegenüber angepasst. Der verwendete Leitfaden umfasst im Wesentlichen zwei große Frageblöcke, wobei der Erste sich mit den Themenbereichen Sammeln, Archivieren, Finden, Bereitstellen und der Zweite mit der Nachnutzung qualitativer Daten beschäftigt. Gerahmt wurden diese beiden Blöcke jeweils von Einstiegs- bzw. Abschluss- und Ausblicksfragen.

Einleitungsfragen:

- Ist die Audioaufzeichnung in Ordnung?
- Können Sie Ihr Projekt und Ihre eigene Aufgabe darin vorstellen?

Block 1: Sammeln, Archivieren, Finden, Bereitstellen:

Allgemeines:

- Welche Arbeitsschritte umfasst die Archivierung?
- Welche Aufgaben übernimmt die Einrichtung?
- Welche Rolle haben die Forschenden bei der (kooperativen) Datenvorbereitung und -verarbeitung?
- Wie unterscheidet sich der Umgang bei verschiedenen Materialarten im Forschungsdatenmanagement? (z. B. analoges, digitales sowie sensibles Material)
- Inwiefern lassen sich generische Arbeitsabläufe entwickeln?
- Wie viel spezifische Anpassung erfordert der Umgang mit den individuellen Daten?
- Wie wird die Finanzierung des Datenmanagements sichergestellt?
- Braucht es einen fachwissenschaftlichen Hintergrund, um in diesem Feld arbeiten zu können?
- Wie sollte das Forschungsdatenmanagement universitär verankert sein?

Erschließung:

- Wie werden die Daten von Ihnen erschlossen?
- Wie wird die Kontextualisierung des Materials sichergestellt?

- Unterscheidet sich die Erschließung je nach Datentyp?
- Basiert die Metadatenvergabe auf kontrolliertem Vokabular/genormten Schlagworten?
- Welche Probleme gibt es im Zusammenhang mit der Erschließung?

Bereitstellung und Auffindbarkeit:

- Wie funktioniert der Prozess der Anonymisierung bei Ihnen?
- Gab es bereits Probleme mit der Anonymisierung?
- Welche Daten werden bei Ihnen im Safe-Room aufbewahrt?
- Wie werden die Kenntnisse systematisch an mögliche Nutzer_innen vermittelt?
- Was umfassen die Schulungen, die Sie anbieten?
- Wie kann damit umgegangen werden, dass Beforschte ebenfalls Zugriff auf die Daten bekommen?
- Woher wissen potenzielle Nutzer_innen, dass eine Quelle existiert?
- In welche Nachweissysteme (Kataloge etc.) sind die Daten eingebettet?
- Wird aktiv Werbung für die von Ihnen verwalteten Daten gemacht?

Auswahl- und Bewertungskriterien:

- Welche Kriterien werden für die Archivierungswürdigkeit angelegt?
- Wer entscheidet über die Archivierung?
- Gibt es eine ethische Verpflichtung bestimmte Sachen aufzubewahren?
- Werden bei der Akquise unmittelbar Personen angesprochen, wenn klar ist, dass sie zu einem interessanten Thema forschen?
- Wird die Auswahl durch die Nähe zu der Fachcommunity unterstützt?
- Wie wird mit analogen Daten umgegangen?

Kontakt und Kooperation mit Wissenschaftler_innen, der Community:

- Mit welchen Problemen kommen Wissenschaftler_innen auf Sie zu?
- Was sind die Bedürfnisse der entsprechenden Community?
- Wie kann eine (vertrauensvolle) Beziehung zur Fachcommunity etabliert werden?
- Werden von außen Anforderungen/Wünsche an Sie herangetragen?

Block 2: Nachnutzung

- Wie würde eine konkrete Nachnutzung in Ihrer Einrichtung ablaufen?

- Welche Daten werden von wem genutzt?
- Was wird mit solchen Daten gemacht?
- Ist die Nachnutzung nur auf die Forschung beschränkt oder gibt es auch Überlegungen für den Bereich Unterricht/Lehre?
- Wo sehen Sie derzeit noch die größten Hindernisse für eine Nachnutzung?
- Gibt es Szenarien, wo eine Archivierung bzw. Nachnutzung nicht möglich ist?

Abschluss- und Ausblicksfragen:

- Wo denken Sie, sollte sich das Forschungsdatenmanagement in den nächsten Jahren hin entwickeln?
- Was werden die größten Herausforderungen sein?
- Welche Themen wurden im Interview noch nicht oder zu wenig behandelt?

8.2) *Transkripte der Leitfadeninterviews zur Studie*

„Spezifika und Herausforderungen qualitativer Daten im Forschungsdatenmanagement - Specifics and challenges regarding research data management with qualitative data“

Interview 1.....	2-24
Datum des Interviews: 27.3.2020	
Dauer des Interviews: 01:05:58	
Interview 2.....	25-27
Datum des Interviews: 27.3.2020	
Dauer des Interviews: 01:15:24	
Interview 3.....	48-52
Datum des Interviews: 9.3.2020	
Dauer des Interviews: ca. 120 Minuten	
Interview 4.....	53-65
Datum des Interviews: 11.3.2020	
Dauer des Interviews: 00:43:09	
Interview 5.....	66-78
Datum des Interviews: 18.3.2020	
Dauer des Interviews: 00:43:28	
Interview 6.....	79-96
Datum des Interviews: 19.3.2020	
Dauer des Interviews: 01:00:14	
Interview 7.....	97-116
Datum des Interviews: 25.3.2020	
Dauer des Interviews: 01:06:14	