

Forschungsdaten

AR 3246

Wjatscheslaw Sterzer und Susanne Kretzer

Archivierungsstrategien anpassen – Herausforderungen und Lösungen für die Archivierung und Sekundärnutzung von ethnologischen Forschungsdaten

Zusammenfassung: Ethnologische Forschungsdaten enthalten in der Regel sensible und schutzwürdige, häufig personenbezogene Daten, die hohe Anforderungen an die Archivierung und Aufbereitung für die Sekundärnutzung stellen. Angemessene und flexible Workflows erlauben eine adäquate Darstellung der Studien in einer sicheren technischen Umgebung und berücksichtigen forschungspraktische, ethische und rechtliche Aspekte. Für den deutschen Raum hat sich das Datenservicezentrum Qualiservice als einzige Möglichkeit erwiesen, um eine adäquate Archivierung für solche Daten zu ermöglichen.

Der Beitrag beschreibt die Anforderungen für die Archivierung von ethnologischen Forschungsdaten und die Workflows, die von Qualiservice für den ethnologischen Bereich adaptiert werden können.

Schlüsselwörter: Ethnologische Forschungsdaten, Archivierung, Qualiservice

Adjusting Archival Strategies – Challenges and Solutions for Archiving and Reusing Ethnological Research Data

Abstract: Ethnological research data contains sensitive and personal information, which imposes very high requirements for its processing, archiving and reuse. Appropriate and flexible workflows enable the adequate description and presentation of studies in a technically secured environment that reflects research practices, ethical and legal aspects within the ethnological context. The data service centre Qualiservice has proven to be the only means in Germany to adequately archive the research data in question. The paper describes the requirements for the archiving of ethnological research data and the workflows that can be adapted for a dedicated ethnological space at Qualiservice.

Keywords: Ethnological research data, archiving, Qualiservice

1 Ausgangslage

Der immer stärker werdende Ruf nach mehr Transparenz und Effizienz in der Forschung hat im nationalen und internationalen Raum zu einem Paradigmenwechsel geführt, der die

Forschungslandschaft nachhaltig prägt. In der Wissensproduktion werden Forschungsdaten von verschiedenen Disziplinen immer häufiger als Supplemente zu den eigentlichen Publikationen nachgefragt bzw. als Rohstoffe für weitere Forschung verwendet. Entsprechend wird von den Förderern die Datenpublikation zunehmend verlangt und die Archivierung von Forschungsdaten vorausgesetzt. Doch während dieser Ruf in den quantitativ forschenden Disziplinen überwiegend stärkeren Zuspruch findet, und entsprechende Institutionen wie z.B. GESIS mit langjähriger Erfahrung zur Verfügung stehen, steckt die Etablierung geeigneter Infrastrukturen für qualitativ forschende Disziplinen noch in den Kinderschuhen. Insbesondere in den Ethnologien¹ stehen die Debatten über Archivierung und Sekundärnutzung von Forschungsdaten aufgrund ihrer hochsensiblen personenbezogenen Natur noch ganz am Anfang.

2 Stärken bündeln – Der Fachinformationsdienst Sozial- und Kulturanthropologie und Qualiservice

Seit 2016 erarbeitet der Fachinformationsdienst Sozial- und Kulturanthropologie (FID SKA) in Zusammenarbeit mit den ethnologischen Fachcommunities fachtypische Anforderungen sowie Praxisempfehlungen für die Archivierung und Sekundärnutzung von Forschungsdaten unter forschungspraktischen, forschungsethischen und gesetzlichen Vorgaben. Um den aktuellen Stand der Forschungspraxis zu ermitteln und die Probleme zu identifizieren, die im Kontext der Archivierung und Sekundärnutzung entstehen können, hat der FID zunächst eine Umfrage² mit Ethnologen durchgeführt. Des Weiteren stand der FID im Austausch mit Repositorien mit ähnlichen Datentypen und Fragestellungen, um geeignete Lösungen und Best Practices der Workflows abzuleiten³. Zurzeit gibt es in Deutschland kein Repository, das die Anforderungen ethnologischer Forschungsdaten fachgerecht abbilden kann. Als anschlussfähig hat sich das Datenservicezentrum (DSZ) Qualiservice⁴ mit seiner organisatorischen, fachlichen und technischen Nähe zu ethnologischen Forschungsdaten in Deutschland herauskristallisiert.

Qualiservice bietet als einziges DSZ die Archivierung und Bereitstellung thematisch ungebundener qualitativer Forschungsdaten aus der Sozialforschung für die Nachnutzung an. Qualiservice verfügt über langjährige Expertise in der Entwicklung angemessener Workflows (Datenaufbereitung, Datennachweissysteme, Datensicherung und Datenbereitstellung), der Beschreibung der Metadaten⁵

¹ Unter „Ethnologien“ bzw. „Sozial- und Kulturanthropologien“ sind die Fachtraditionen Völkerkunde (Ethnologie) und Volkskunde (Europäische Ethnologie) im deutschsprachigen Raum zu verstehen.

² Imeri und Danciu (2017).

³ Harbeck et al. (2018).

⁴ www.Qualiservice.org.

⁵ Betancort und Haake (2014).

sowie Vorschläge für die Kontextualisierung der Studien.⁶ Qualiservice erarbeitet Lösungen für die komplexen rechtlichen Fragen der Sekundärnutzung, insbesondere Anonymisierungsstrategien⁷ im Sinne des Datenschutzes.⁸ Die Arbeit von Qualiservice ist von dem Gedanken getragen, dass der Umgang mit sensiblen qualitativen Daten angemessene Workflows und Standards erfordert, die im Diskurs mit der Scientific Community und ihren Fachgesellschaften erörtert werden.⁹ Für den regen Austausch mit der (sozialwissenschaftlichen) Fachöffentlichkeit organisiert Qualiservice Workshops, beteiligt sich an Veranstaltungen der Fachöffentlichkeit sowie an bestehenden Netzwerken. Die Expertise von Qualiservice wird inzwischen in den verschiedensten sozial- und kulturwissenschaftlichen Disziplinen nachgefragt.

Im Rahmen eines Verbunds mit dem interdisziplinär, mehrfach zertifizierten und nachhaltig arbeitenden Datapublisher PANGAEA,¹⁰ wird Qualiservice derzeit DFG-gefördert zu einem nationalen Regelarchiv für qualitative Forschungsdaten mit dem bisherigen Schwerpunkt auf Interviewdaten ausgebaut. PANGAEA verfügt über eine Infrastruktur, die eine Adaption an die Anforderungen der Archivierung und Bereitstellung kaum standardisierter qualitativer Forschungsdaten ermöglicht und auch die Archivierung weiterer Datenformate, etwa audiovisueller Daten, erlaubt. Weitere Partner sind das GESIS Leibniz-Institut für Sozialwissenschaften Köln für die verteilte Archivierung von Mixed-Methods-Studien und die Staats- und Universitätsbibliothek Bremen, die ihre Expertise bei der Entwicklung von Metadaten für die Darstellung qualitativer Daten zur Verfügung stellt und außerdem die Studienreports der archivierten Forschungsprojekte publiziert.

Geplant ist, in einer beantragten zweiten Förderphase des FID SKA (2019–2021) in Kooperation mit Qualiservice einen an die Anforderungen an ethnologische Daten spezifizierten Bereich im DSZ zu etablieren und damit die Archivierung und Sekundärnutzung ethnologischer Daten zu befördern.

3 Ethnologische Forschungsdaten und Praxis im Kontext

Ethnologische Forschung beruht auf der Interaktion der Forschenden mit Menschen, dem Betreten des persönlichen Lebensraumes der Protagonisten über längere Zeiträume, in denen Ethnologen tiefgreifende Einblicke in soziale Phänomene und Strukturen etwa durch teilnehmende Beobachtung erhalten. Durch multimethodische Erhebungsstrategien entstehen überwiegend hochsensible, miteinander in Beziehung stehende heterogene Forschungsdaten, die Domänen wie z. B. religiöse Praktiken, politische und sexuelle Orientierung oder Gesundheit im Detail beschreiben und daher aus

⁶ Kretzer (2013a), Medjedović und Witzel (2010), Hollstein (2011), http://qualiservice.org/fileadmin/text/Leitfaden_Studienreport_Kontexte.pdf.

⁷ Kretzer und Witzel (2013), Kretzer (2013b).

⁸ Gebel et al. (2015).

⁹ Kretzer (2013a).

¹⁰ <https://www.pangaea.de>.

ethischen und rechtlichen Gründen fachgerecht geschützt werden müssen. Das Vertrauensverhältnis beruht auf dem Versprechen, involvierte Personen und Personengruppen vor Schaden wie Diskriminierung, Verfolgung oder jeglicher Beeinträchtigung persönlicher Beziehungen zu bewahren. Die kritische Auseinandersetzung mit möglichen Risiken sowie deren Minimierung ist für Forschende als Verwahrer potentiell kritischer Daten der ethische Grundpfeiler ethnologischer Forschung.¹¹ Anders als bei quantitativer Forschung, bei der die Erhebungsmethoden meist standardisiert sind, lassen sich Forschungsdaten aus der qualitativen Sozialforschung aufgrund der vielfältigen und komplexen Erhebungsmethoden nicht so einfach kategorisieren.¹² Fachtypisch ist die Planung der Forschungsprozesse aufgrund der Unvorhersehbarkeit im Feld nur bedingt möglich, was zur Folge hat, dass Forschungsmethoden nicht nur wenig standardisiert sind, sondern in manchen Situationen ad hoc neu konzipiert werden müssen.

Aus diesen Dynamiken resultieren mit Blick auf die entstehenden Forschungsdaten ebenso dynamische Anforderungen an die Archivierung und Sekundärnutzung. Wir sind zu dem Schluss gekommen, dass die erfolgreiche Umsetzung davon motiviert werden muss, dass:

- a. Forschungsdaten und ihre Beziehungen zueinander authentisch erhalten bleiben,
- b. der Daten- und Persönlichkeitsschutz involvierter Personen gewährleistet wird und
- c. die forschungsethischen Grundsätze der Ethnologien adäquat abgebildet werden.

4 Komplexe Daten - Heterogenität und Relationen

Fachtypisch entsteht in der ethnografischen Feldforschung eine Vielzahl an diversen Forschungsdaten und Formaten, die die Komplexität sozialer Situationen aus mehreren Blickwinkeln einfangen und auf verschiedenen Wegen beschreiben.

¹¹ Breidenstein et al. (2015).

¹² Huschka und Oellers (2013) 11.

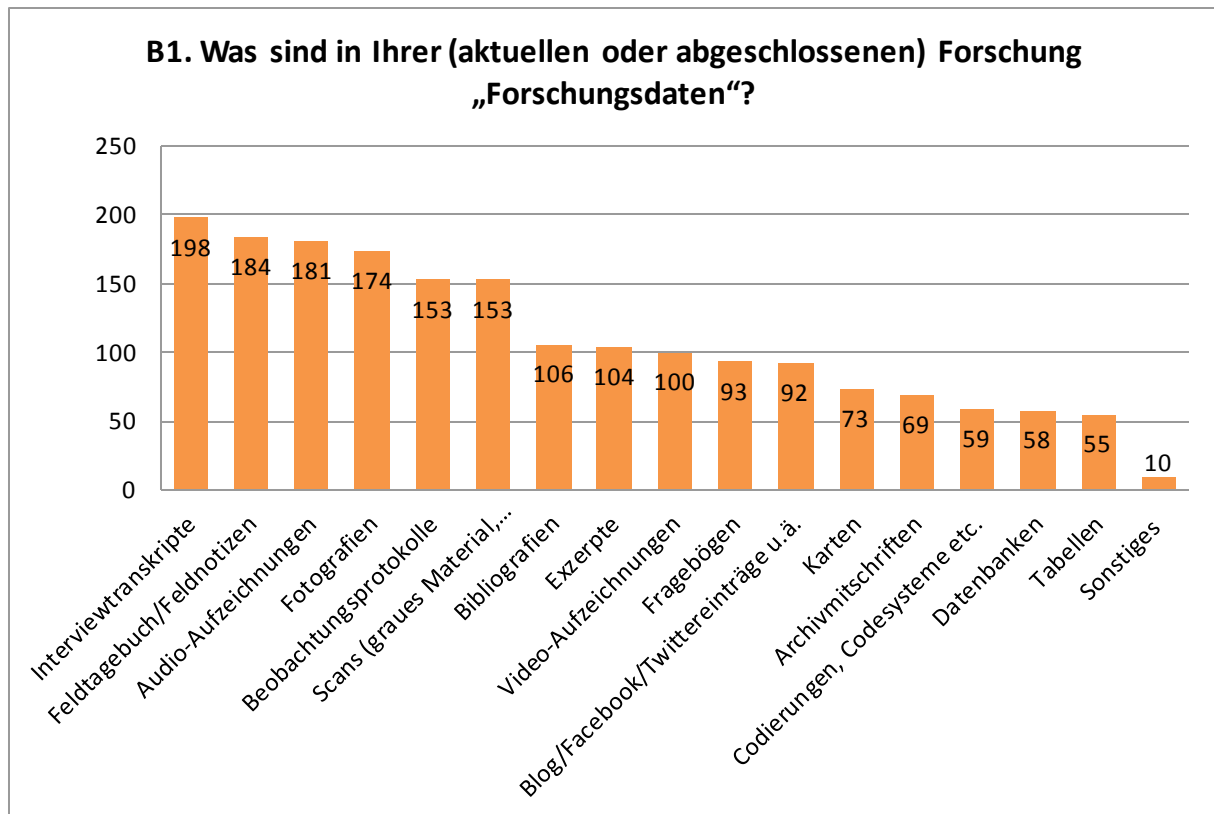


Abb. 1: Was sind in Ihrer (aktuellen oder abgeschlossenen) Forschung Forschungsdaten? (N=251)¹³

Es ist gängige Praxis, sich referenzierende Fotoaufnahmen, Notizen und Memos etc. mit anderen Primärdaten wie z.B. einem Interviewtranskript zu verknüpfen, um ein klareres Bild einer Situation zu erhalten. Dabei ist die entstehende Verknüpfung dieser Datensets für den Kontext der jeweiligen Situation bzw. für das Gesamtverständnis unabdingbar. Für die Archivierung bedeutet das, dass nicht nur die eigentlichen Primärdaten als wissenstragende Objekte relevant sind, sondern Verknüpfungen und Relationen der jeweiligen Primärdaten das eigentliche Forschungsdatenset vervollständigen. Sollten Forschungsdaten aber isoliert von ihren Kontexten und Relationen archiviert werden, würden die Transparenz und auch die Nachvollziehbarkeit immens leiden.

Um die Forschungspraxis nicht durch technische Vorgaben für die Datenübergabe zu beeinträchtigen, stellt die Berücksichtigung von heterogenen Datenformaten eine weitere Schlüsselanforderung an das DSZ dar. Weil die Forschungsmethoden heterogen sind, gibt es per se keine Standardsoftware, mit der Daten erhoben werden. Textverarbeitungsprogramme und Aufnahmegeräte fallen je nach Equipment und Präferenz der Primärforschenden sehr unterschiedlich aus. Nicht selten wird auf Werkzeuge zurückgegriffen, die im Augenblick für die Erschließung eines bestimmten Forschungsaspekts am sinnvollsten erscheinen. Darüber hinaus findet spezialisierte Software in der Analyse Verwendung, um erhobene Daten mit Analysewerkzeugen wie z.B. MAXQDA zu organisieren und zu bearbeiten. Zu bedenken gilt allerdings, dass kommerzielle Softwarelösungen

¹³ Imeri und Danciu (2017) 10.

oft exotische proprietäre Formate generieren, die sich auch nur mit entsprechend gültiger Lizenz wieder öffnen und bearbeiten lassen.¹⁴ Für den Fall, dass solche Formate archiviert werden sollten, könnten zusätzlich zum exotischen Format auf dem Wege der Exportfunktion einzelne Datensätze in *gängige*¹⁵ Formate exportiert werden.

Weil der Versuch, diese Heterogenität auf eine überschaubare Liste von Formaten zu reduzieren, Einschränkungen für die Datenerhebungsmethoden und Forschungspraxis der Forschenden nach sich ziehen könnte, sind wir der Meinung, dass die Auswahl geeigneter Formate für die Archivierung nicht nur technisch motiviert sein darf, sondern ebenso die Natur der ethnologischen Forschung im Archivierungsworkflow reflektiert werden muss. Daher empfehlen wir, dass alle gängigen Formate, die im Forschungsalltag durch diverse Software und Hardware entstehen können, Berücksichtigung finden.

5 Rechtliche und forschungsethische Herausforderungen

Die Aufbereitung und Verwaltung von sensiblen personenbezogenen Forschungsdaten stellt eine besondere Herausforderung dar, der auf technischer Ebene nur mit erheblichem Aufwand begegnet werden kann. Zunächst erfordern sensible Forschungsdaten ein sehr hohes Maß an Kontrolle und Vorkehrungen, um den strikten Persönlichkeits- und Datenschutz der beteiligten Protagonisten zu gewährleisten. Ähnlich wie das in anderen Datenrepositorien bereits Praxis ist, ist auch mit Blick auf ethnologische Forschungsdaten ein gestaffeltes Rechtemanagement für die differenzierte Abbildung der Zugangsrechte sinnvoll und angemessen. Die Differenzierung sollte sich an das Datenschutzrisiko des jeweiligen Materials lehnen und die Zugriffe individuell regulieren. So wären z. B. Metadaten freisuchbar, jedoch anonymisierte Interviews nur mit einem Nutzungsantrag zugänglich. Bild-, Audio- und Videomaterial würde, aufgrund des höheren Datenschutzrisikos, eine höhere Schutz- und Zugriffsstufe zugeordnet werden.¹⁶

Es stellt sich zudem die Frage, inwieweit die Anonymisierung bestimmter Forschungsdaten das Nachnutzungspotential beeinträchtigen könnte. Während durch Anonymisierung und Pseudonymisierung der Re-Identifikation der Protagonisten in Interviewtranskripten vorgebeugt werden kann, wird es bei Bild-, Audio- und Videodaten umso schwieriger, bzw. kann die Sinnhaftigkeit der Anonymisierung dieser Daten in Frage gestellt werden. Zum Beispiel wäre es

¹⁴ In MAXQDA als Projekt abgespeicherte Datensets werden im proprietären **.mx** Format abgespeichert, welches nicht kompatibel mit anderer Software ist. Die weitere Bearbeitung dieser Datensets setzt daher zwingend die Verwendung der Software voraus, was mit hohen Lizenzkosten verbunden sein kann.

¹⁵ Als *gängig* soll an dieser Stelle der *Verbreitungsgrad* und die *Kompatibilität* als Referenz gelten. So sind z. B. Microsofts Office Formate wie MS Word **docx** und MS Excel **xlsx** zwar proprietär, jedoch ist der Verbreitungsgrad enorm und der Grad der Kompatibilität ist hoch, sodass eine docx und xls auch mit Open Source Software wie Open Office problemlos geöffnet und bearbeitet werden kann.

¹⁶ Harbeck et al. (2018).

durchaus möglich, Bild-, Audio- und Videodaten durch Verpixelung der Gesichter oder Verzerrung der Stimmen mit Open Source Software (z.B. GIMP, Audacity, OpenShot) zu anonymisieren. Jedoch ergeben sich zwei große Problemfelder für die Sekundärnutzung mit diesem Ansatz.

1. Die Anonymisierung von Bild-, Audio- und Videomaterial würde die Interpretierbarkeit und damit das Nachnutzungspotential des Materials immens beeinträchtigen oder es sogar unbrauchbar machen.
2. Die Anonymisierung kann mit geeigneter Software ggf. auch wieder rückgängig gemacht werden.

Die größte Herausforderung besteht darin, die Balance zwischen Datenschutz und Nachnutzungspotential im Rahmen gesetzlicher und ethischer Vorgaben zu finden. Dies könnte insbesondere für audiovisuelle Forschungsdaten mit strengeren Zugriffsrechten reguliert werden.

6 Archivierung qualitativer Studien forschungsfreundlich gestalten

Die vorangegangenen Ausführungen bezeichnen die Anforderungen für Datenservicezentren, die ethnologische qualitative Forschungsdaten archivieren und für die Nachnutzung aufbereiten wollen. Qualiservice hat Workflows konzipiert, die qualitatives Forschungsvorgehen abbilden, ohne den Wissenschaftlern starre Archivierungsgerüste aufzuzwingen. Diese flexiblen Workflows können auf die Anforderungen aus den Ethnologien angepasst und erweitert werden. Gleichwohl sollte die Einhaltung von Qualitätskriterien und angemessenen Standards in der Datenaufbereitung es Sekundärnutzern ermöglichen, Forschungsdaten für ihr Forschungsanliegen sicher zu beurteilen (Data Fit).¹⁷ Eine kooperative Datenvorbereitung von Qualiservice und Primärforschern gewährleistet, dass vorhandene Workflows und Qualitätskriterien für die Sekundärdaten aus der ethnologischen Forschungspraxis heraus entwickelt und diese Erfahrungen in den ethnologischen Communities und ihren Fachgesellschaften spezifiziert werden können.

Die Begleitung der Datenvorbereitung findet im persönlichen Kontakt in der individuellen Beratung von Wissenschaftlern und in Workshops statt. Ab 2019 veranstaltet Qualiservice Workshops, die interessierten Datengebern die Möglichkeit bieten, ihr Datenmanagement und ihre Forschungsdokumentation hinsichtlich der Sekundärnutzung zu planen, notwendige Kontextualisierungen auf Studien- und Primärdatenebene sowie ihr Vorgehen bei der Anonymisierung zu systematisieren. Der regelmäßige Kontakt zwischen Qualiservice und Forschern erlaubt es, rechtliche und forschungsethische Überlegungen zu Archivierung und Sekundärnutzung der einzelnen Studien fortlaufend zu reflektieren und notwendig gewordene Modifikationen im Forschungsprozess bspw. bei der Datenerhebung in der Datenvorbereitung zu berücksichtigen. Forscher können zukünftig die direkte Kommunikation mit Qualiservice auch auf der DSZ-eigenen

¹⁷ Medjedović und Witzel (2010) 31ff.

Website über ein Ticketsystem¹⁸ auslösen. Der Austausch mit Qualiservice findet in einer sicheren dialogischen Verbindung statt und wird fortlaufend dokumentiert. Die kooperative Datenvorbereitung schon in der Primärforschung unterstützt eine zeitnahe Studienübergabe nach dem Projektabschluss an Qualiservice. Die Datenübergabe erfolgt über einen Upload-Space, der ebenfalls zukünftig über die Qualiservice-Website zu finden ist. Qualiservice bietet Forschern an, schon während des Forschungsprozesses Daten zu hinterlegen, dies kann bspw. eine Entlastung der Forscher während langer Forschungsreisen bewirken.

Für die Übergabe der Daten schließt Qualiservice eine Vereinbarung mit Datengebern ab. Die Vereinbarung bietet Forschenden verschiedene Optionen, um Bedingungen für die Sekundärnutzung ihrer Studien niederzulegen. Beispiele sind ein zeitliches Embargo oder der Ausschluss bestimmter Nutzungszwecke wie etwa die Verwendung des Datensatzes in der Lehre. Gemeinsam mit den Forschern wird eine Sicherheitskategorie für die Forschungsdaten festgelegt. So ist es z.B. möglich, dass Primärdaten nur im Safe Center eingesehen werden dürfen. Diverse Formate der Primärdaten, etwa Transkripte und audiovisuelle Daten können unterschiedlichen Sicherheitskategorien zugordnet werden, bleiben dabei aber in ihrem Forschungszusammenhang organisiert. Die Sekundärnutzung der Studien ist immer mit einer Nutzungsvereinbarung verbunden.

Obligatorisch für die Archivierung bei Qualiservice ist die Übergabe der Einverständniserklärung der Forschungsteilnehmer für die Archivierung und Sekundärnutzung durch den Datengeber. Allerdings steht Qualiservice im Austausch mit Datenschützern, um eine Lösung für Forschungssituationen zu finden, in denen Einverständniserklärungen nicht konform zum neuen Datenschutzrecht zu erlangen sind. Dies ist insbesondere für politisch brisante Forschungen, etwa über Rechtsextremismus, erforderlich, aber auch für ethnologische Studien, etwa in Milieus am Rand der Legalität oder wenn beforschte Gruppen von Stigmatisierung und Repressionen betroffen sind¹⁹ und sollte in einem ethischen Diskurs in der Scientific Communities erörtert werden.

Die Kuration der Studien findet im DSZ-eigenen Safe Center statt. Dort werden direkte Identifikationsdaten wie bspw. Kontaktdaten aus den Forschungsdaten entfernt, auf einem separaten Rechner und örtlich getrennt gespeichert. Im Safe Center wird anschließend u. a. die Anonymisierung/Pseudonymisierung der Daten geprüft und zu abgabefähigen Datensätzen für die Sekundärnutzung einerseits und zu Datensätzen für die Langzeitarchivierung andererseits aufbereitet. Die Studien bekommen mit der Veröffentlichung einen DOI zur sicheren Findbarkeit und sind zitierbar. Die Datengeber können den aufbereiteten Datensatz mit einem proof read begutachten.

¹⁸ Dieser Service wird zum Jahresanfang 2019 in Betrieb genommen.

¹⁹ Imeri (2018) 5-6.

Bei Vorliegen verschiedener Datentypen (Interview, Bild- und Audioformate, Beobachtung), etwa durch die Kombination verschiedener Erhebungsmethoden in einer Studie, werden die diversen Daten unterschiedlich für die Archivierung und Sekundärnutzung aufbereitet. Dies liegt z.B. in unterschiedlichen Anforderungen an die Anonymisierung begründet bzw. an deren unterschiedlicher Auswirkung auf das Sekundärnutzungspotenzial der Primärdaten. Entsprechend der Datenvielfalt müssen sowohl unterschiedliche Kurations- und Sicherheitsmaßnahmen als auch Nachnutzungsoptionen eingerichtet und für die langfristige Nutzbarkeit der Daten unterschiedliche Erhaltungsmaßnahmen in der Langzeitarchivierung ergriffen werden.

Qualiservice arbeitet fortlaufend an einem System der flexiblen Anonymisierung für ein erhöhtes sekundäres Nutzungspotenzial der textgebundenen Primärdaten. Um das Legen falscher Fährten für die spätere Dateninterpretation zu vermeiden, favorisiert Qualiservice die Anonymisierung durch das Abstrahieren der konkreten Information in die relevante sozialwissenschaftliche Information (bspw. wird ein Herzinfarkt zur koronaren Herzkrankheit). Einzelne konkrete Informationen, die für eine bestimmte Forschung, bspw. eine genaue Berufsbezeichnung für eine berufspädagogische Nachnutzung, benötigt werden, können gegeben werden. So ergibt sich je nach Forschungsanliegen jeweils eine andere Kombination aus anonymisierten und konkreten Informationen, eine erschwerte Re-identifizierung von Forschungspartnern bleibt gewährleistet.

7 Metadaten

Die generierten Metadaten zur Studie²⁰ werden in englischer und deutscher Sprache veröffentlicht und mit dem Studienreport, der keine sensiblen Daten beinhaltet, verlinkt. So können Datensuchende unmittelbar einen Eindruck über die Eignung einer Studie für ihre Sekundärforschung gewinnen. Die Staats- und Universitätsbibliothek veröffentlicht die Studienreports über den DINI-zertifizierten Dokumentenserver E-LIB mit einer URN-Registrierung. In Kooperation mit der DDI Working Group on Qualitative Data²¹ erarbeitete Qualiservice ein Metadatenschema für qualitative Daten. In diesem – ursprünglich für quantitative Daten entwickelten – Schema wurden Kategorien geöffnet, um die Darstellung bspw. von qualitativen Methoden beschreiben zu können. Das erarbeitete Metadatenschema ist anschlussfähig an das von PANGAEA genutzte panFMP²² und weitere Metadatenschemata.

Metadaten zu den Primärforschungsdaten (Mikrometadaten) werden nicht in Katalogen oder in Portalen publiziert, da sie über die Kombination der Informationen die Anonymität von Forschungspartnern gefährden könnten. Die Primärforschungsdaten werden über Uniform Resource

²⁰ Das Metadatenschema wird zurzeit aktualisiert und anschließend auf www.qualiservice.org veröffentlicht.

²¹ Hoyle et al. (2013).

²² <http://panfmp.org>.

Identifizier (URIs) organisiert. Datensuchende können, wenn sie potenziell nutzbare Studien für ihr Forschungsinteresse gefunden haben, gezielt in den Mikrometadaten Qualiservice-intern nach passenden Datensätzen suchen.

8 Der Informatikpartner PANGAEA23

PANGAEA ist Mitglied des “International Council for Science World Data System (ICSU-WDS)”²⁴ und ist als “World Data Center (WDC)” zertifiziert und akkreditiert. Das Datenzentrum wird entsprechend den FAIR²⁵ Prinzipien betrieben. Die Metadaten der archivierten Studien werden in zahlreichen internationalen, teils auch Community-spezifischen Datenportalen veröffentlicht. Zu archivierende Daten werden durch Domainexperten – bei Qualiservice durch sozialwissenschaftliches Fachpersonal in Kooperation mit den beteiligten Forschern – qualitätsgeprüft sowie strukturell und semantisch über das “Editorial System” harmonisiert. Grundlage dafür sind Inhalts-Standards und relevante Terminologien der jeweiligen Communities.²⁶ Für qualitative Daten können an die Besonderheiten qualitativen Forschens angepasste Inhalts-Standards im Datenbanksystem für die authentische Studiendarstellung und Terminologien (kontrolliertes Vokabular) im Austausch mit Forschenden und aus der Archivierungspraxis heraus entwickelt bzw. implementiert werden (z.B. die Gemeinsame Normdatei GND). Insbesondere können sensible Daten in flexibler Weise geschützt werden. Unterschiedliche Nutzungszugänge können durch Kuratoren und Datengeber definiert und mit einer entsprechenden Nutzerregistrierung freigegeben werden. Nutzer können ihre Authentifizierung über ihre ORCID²⁷ vornehmen. Sensible Daten, die Qualiservice zugeordnet sind, sind nicht frei zugänglich. Als multidisziplinäres Informationssystem werden seit ca. 25 Jahren schwerpunktmäßig geo- und biowissenschaftlichen Daten archiviert und publiziert. Durch die Nutzung vielfältiger Datenformate in diesen Disziplinen ist PANGAEA in der Lage, unterschiedlichste Formate für Text-, Audiodateien und audiovisuellen Daten speichern und darstellen zu können. Für die Ethnologien bringt diese Vielfältigkeit den Vorteil mit sich, bspw. im Feld nach forschungspraktischen Aspekten Formate auswählen zu können.

²³ Kretzer und Diepenbroek (2018).

²⁴ <https://www.icsu-wds.org/>

²⁵ <https://www.force11.org/fairprinciples>

²⁶ Diepenbroek et al. (2017).

²⁷ <https://orcid.org/> - ORCID ermöglicht u.a. die eindeutige Zuordnung von Nutzern zu Publikationen von Daten und Literatur.

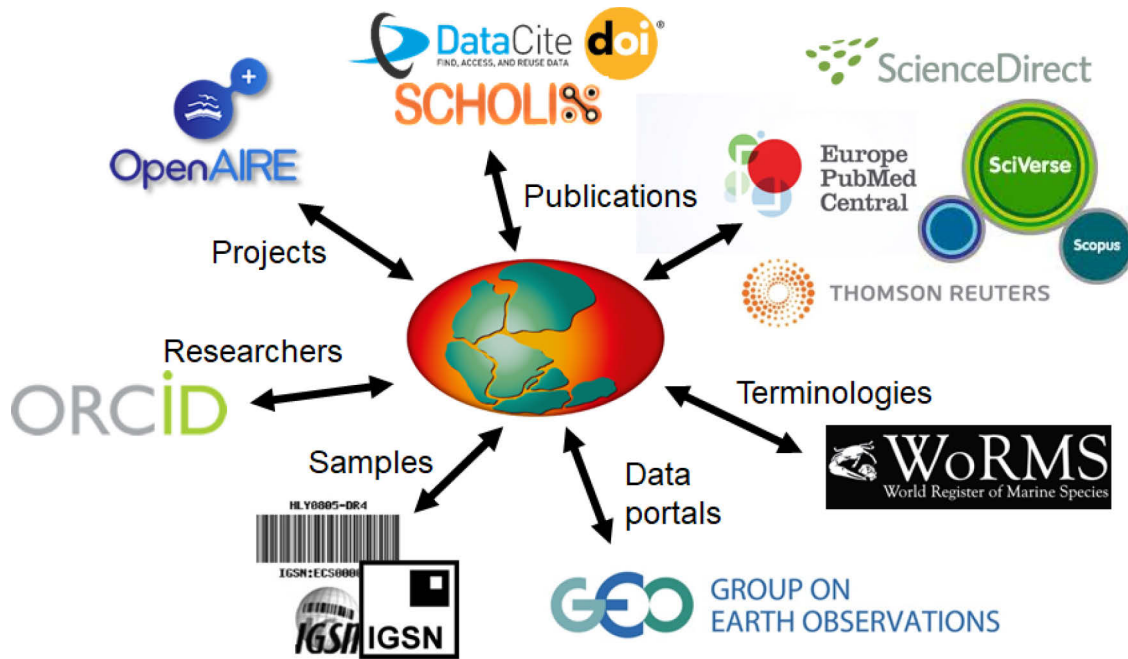


Abb. 2: Vereinfachte Darstellung der Interoperabilität von PANGAEA

PANGAEA ermöglicht eine weitgehende Interoperabilität der archivierten Datenbestände (Abb. 2) als zusätzliche Services. Hervorzuheben ist das „cross-linking“ zwischen Literatur und Daten. Verlage können über den mit maßgeblicher Beteiligung von PANGAEA aufgebauten Dienst SCHOLIX²⁸ Daten referenzieren, die mit Zeitschriftenaufsätzen relationiert sind.

9 Schlussbemerkung

Die Kooperation des FIDSKA, der ethnologischen Communities und Qualiservice mit seinen Projektpartnern bietet gute Voraussetzungen, um in einem gemeinsamen Klärungsprozess eine flexible und nachhaltige, an die besonderen Anforderungen der Ethnologien ausgerichtete Praxis der Archivierung, Aufbereitung und Bereitstellung für die Sekundärnutzung von Forschungsdaten zu entwickeln und umzusetzen. In naher Zukunft ist es geplant, von Qualiservice in realen Forschungsprozessen schon entwickelte Workflows zu erproben und an die ethnologischen Bedarfe anzupassen.

Literaturverzeichnis

Betancort Cabrera, Noemi; Haake, Elmar (2014): Das QualiService-Metadatenschema, Version 1.1. QualiService Technical Reports, 2014/01. Verfügbar unter <http://nbn-resolving.de/urn:nbn:de:gbv:46-00103643-13>.

Breidenstein, Georg; Hirschhauer, Stefan; Kalthoff, Herbert; Nieswand, Boris (2015): Ethnografie: Die Praxis der Feldforschung. 2. Aufl. Konstanz: UVK Verlagsgesellschaft mbh.

²⁸ <http://www.scholix.org/>

- Diepenbroek, Michael; Schindler, Uwe; Huber, Robert; Pesant, Stéphane; Stocker, Markus; Felden, Janine; Buss, Melanie; Weinrebe, Matthias (2017): Terminology supported archiving and publication of environmental science data in PANGAEA. In: *Journal of Biotechnology*, 261, 177–86. Verfügbar unter <https://doi.org/10.1016/j.jbiotec.2017.07.016>.
- Gebel, Tobias; Grenzer, Matthias; Kreuzsch, Julia; Liebig, Stefan; Schuster, Heidi; Tschewinka, Ralf; Watteler, Oliver; Witzel, Andreas (2015): Verboten ist, was nicht ausdrücklich erlaubt ist: Datenschutz in qualitativen Interviews. In: *Forum Qualitative Sozialforschung*, 16(2), Art. 27. Verfügbar unter <http://www.qualitative-research.net/index.php/fqs/article/view/2266/3822>.
- Harbeck, Matthias; Imeri, Sabine; Sterzer, Wjatscheslaw (2018): Feldnotizen und Videomitschnitte. Zum Forschungsdatenmanagement qualitativer Daten am Beispiel der ethnologischen Fächer. In: *Das offene Bibliotheksjournal*, 123–41. Verfügbar unter <https://doi.org/10.5282/obib/2018H2S123-141>.
- Hollstein, Betina (2011): Qualitative Approaches. In: *Sage Handbook of Social Network Analysis*, hg. v. J. Scott und P.J. Carrington, London/New Delhi: Sage, 404–17.
- Huschka, Denis; Oellers, Claudia (2013): Einführung: Warum qualitative Daten und ihre Sekundäranalyse wichtig sind. In: *Forschungsinfrastrukturen für die qualitative Sozialforschung*, 9–16. Verfügbar unter https://www.ratswd.de/dl/downloads/forschungsinfrastrukturen_qualitative_sozialforschung.pdf.
- Hoyle, Larry and the DDI Qualitative Data Working Group: Corti, Louise; Gregory, Arofan; Martinez, Agustina; Wackerow, Joachim; Alvar, Eirik; Betancort Cabrera, Noemi; Gallagher, Damien; Gebel, Tobias; Hautamaki, Jani; Kuula, Arja; McEachern, Steve; Zuell, Cornelia (2013): A Qualitative Data Model für DDI. Verfügbar unter <https://www.ddialliance.org/sites/default/files/AQualitativeDataModelForDDI.pdf>.
- Imeri, Sabine; Danciu, Ida (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. Verfügbar unter <http://www.evifa.de/cms/ueber-evifa/forschungsdatenmanagement>.
- Imeri, Sabine (2018): Archivierung und Verantwortung. Zum Stand der Debatte über den Umgang mit Forschungsdaten in den ethnologischen Fächern. In: RatSWD: *Tagungsband zum Workshop „Archivierung und Zugang zu qualitativen Daten“*, OutputSeries. Verfügbar unter <https://www.ratswd.de/publikationen/output>.
- Kretzer, Susanne (2013a): Infrastruktur für qualitative Forschungsprimärdaten – Zum Stand des Aufbaus eines Datenmanagementsystems von QualiService. In: *Forschungsinfrastrukturen für die qualitative Sozialforschung*. hg. v. Denis Huschka, Hubert Knoblauch, Claudia Oellers und Heike Solga. Berlin: Scivero, 93–110. Verfügbar unter <https://www.ratswd.de/publikationen/forschungsinfrastrukturen-qualitative-sozialforschung>.
- Kretzer, Susanne (2013b): Arbeitspapier zur Konzeptentwicklung der Anonymisierungs-/Pseudonymisierung in Qualiservice. Verfügbar unter <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-47605-2>.
- Kretzer, Susanne; Diepenbroek, Michael (2018): Flexible Strategien für eine forschungsfreundliche Archivierung und Nachnutzung von qualitativen Forschungsdaten. In: RatSWD: *Tagungsband zum Workshop „Archivierung und Zugang zu qualitativen Daten“*, OutputSeries. Verfügbar unter <https://www.ratswd.de/publikationen/output>.
- Kretzer, Susanne; Witzel, Andreas (2013): Datenarchivierung. In: Sektion Biographieforschung in der Deutschen Gesellschaft für Soziologie: Newsletter/Rundbrief 65/Dezember 2013, 27–31.
- Medjedović, Irena; Witzel, Andreas (unter Mitarbeit von Mauer, Reiner und Watteler, Oliver) (2010): Wiederverwendung von qualitativen Daten. Archivierung und Sekundärnutzung qualitativer Interviewtranskripte. Wiesbaden: VS-Verlag für Sozialwissenschaften.
- Sterzer, Wjatscheslaw; Imeri, Sabine; Harbeck, Matthias (2018): Zugriff auf ethnologische Forschungsdaten: Anforderungen und Lösungen. Verfügbar unter <https://opus4.kobv.de/opus4-bib-info/frontdoor/index/index/docId/15719>.



Wjatscheslaw Sterzer

Fachinformationsdienst Sozial- und Kulturanthropologie
Bereich Forschungsdatenmanagement
Humboldt-Universität zu Berlin
Universitätsbibliothek
Unter den Linden 6
D-10099 Berlin
wjatscheslaw.sterzer.1@ub.hu-berlin.de



Susanne Kretzer

Universität Bremen
SOCIUM Forschungszentrum Ungleichheit und Sozialpolitik
Postfach 33 04 40
D-28334 Bremen
skretzer@uni-bremen.de