

Positionspapier zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten¹

Über Chancen und Risiken der digitalen Archivierung von Forschungsdaten wird derzeit in vielen Fächern diskutiert, ebenso werden Strategien entwickelt, Datenmanagement und Anforderungen an künftige Datenarchive umzusetzen, die den je spezifischen Fachkulturen entsprechen. Forciert wird die Debatte durch den Medienwandel selbst, der Forschungspraxis verändert, zu einem guten Teil aber auch durch Forderungen, mit öffentlichen Mitteln erhobene Forschungsdaten langfristig zu sichern und für die weitere Verwendung zugänglich zu machen. Die *Deutsche Gesellschaft für Volkskunde* (im Folgenden mit *dgv* abgekürzt) folgt mit vorliegendem Positionspapier dem Wunsch der DFG, den Umgang mit Forschungsdaten in den Fächern selbst zu reflektieren sowie „angemessene Regularien zur disziplinspezifischen Nutzung und ggf. offenen Bereitstellung von Forschungsdaten zu entwickeln.“² Die Stellungnahme erfolgt zu diesem Zeitpunkt angesichts der Dringlichkeit, auch wenn viele der technischen Aspekte und rechtlichen Regularien in den von Datenmanagementprozessen berührten Bereichen noch nicht ausreichend entwickelt sind. Entsprechend bilden die hier formulierten Positionen einen Stand der Debatte im Fachzusammenhang Europäische Ethnologie³ ab. Sie sind angesichts dynamischer Prozesse der Fachentwicklung, der Technologieentwicklung sowie wissenschaftspolitischer Entscheidungen vorläufig und werden fortlaufend aktualisiert werden.

Die *dgv* erwartet, dass Förderorganisationen die hier formulierten Positionen zum Umgang mit Forschungsdaten bei der Vergabe von Fördermitteln und der Begutachtung von Projektberichten berücksichtigen.

1. Forschungspraxis/ Forschungsdaten in der Europäischen Ethnologie

Die Europäische Ethnologie ist durch eine erhebliche Diversität der Zugänge und methodische Vielfalt gekennzeichnet. Überwiegend wird mit qualitativen Verfahren gearbeitet, im Rahmen ethnografischer Verfahren der Feldforschung und teilnehmenden Beobachtung, verschiedener Formen von Interviews, Archivstudien oder Diskursanalysen.

Insbesondere ethnografische Forschung – die hier schwerpunktmäßig adressiert wird – wird als offener Prozess konzipiert, der situations- und beobachterabhängig verläuft. Gesprächspartner*innen werden dabei nicht als Studienteilnehmer*innen konzeptualisiert und selten als Sample rekrutiert, sie werden vielmehr als mit Rechten ausgestattete Mitglieder eines sozialen Zusammenhangs angesehen, zu dem sie Forschenden Zugang gewähren. Forschende sind in der Regel auf Zustimmung und Kooperation der Erforschten angewiesen, auch werden zunehmend kollaborative Formen der Wissensproduktion und -repräsentation entwickelt. Entsprechend wird die Beziehung zwischen Forschenden und Erforschten als ein beiderseitiges Vertrauensverhältnis verstanden, das die fragile Basis vieler Feldforschungen darstellt.

In der Europäischen Ethnologie sind – ähnlich wie in anderen Sozial-/Geisteswissenschaften – Daten untrennbar an theoretisch begründete Vorannahmen und analytische Perspektivierungen gebunden, die schon in die Konstitution eines Forschungsfeldes eingehen. Entsprechend ist ein Datenbegriff, wie er in Konzepten von „Rohdaten“ oder Primärdaten häufig Anwendung findet (dekontextualisiert, außerhalb von Interpretationssystemen existierend) für Daten aus qualitativer Forschung nicht zutreffend. Der ethnografisch erforschte Ausschnitt des Alltagslebens umfasst vielmehr eine oft unübersehbare Menge an potenziell relevanten Beobachtungsdaten und Inhalten, aus denen Forschende auf der Grundlage theoriegeleiteter Fragestellung und im Rahmen fachspezifischer Methodenstandards eine Auswahl treffen, die die eigentlichen Forschungsdaten generiert.

Die in Forschungen der Europäischen Ethnologie erzeugten Daten sind in der Regel heterogen, wenig standardisiert und multimodal (Text, Bild, Audio, Video, Karten, Webseiten etc.). Aufgrund ihrer Kontextgebundenheit sind sie überwiegend unikal. Der Aufwand, der durch die Aufarbeitung von ethnografischen – oder allgemeiner qualitativen – Daten für Archivierung und Nachnutzung entsteht, ist dementsprechend hoch und komplex. Bisher stehen überdies keine nicht-proprietären, also: *open-source tools* zur Verfügung, die vor allem den ethnografischen ‚Daten-Workflow‘ und die Archivierung originär und spezifisch unterstützen könnten.

Die Nachvollziehbarkeit von aus ethnografischer Forschung gewonnenen Erkenntnissen wird durch die Tatsache gerahmt, dass diese Erkenntnisse immer – notwendig kontingente – Interpretationen sozialer Wirklichkeiten sind. Die Europäische Ethnologie verfügt deshalb über etablierte und wirkungsvolle Verfahren der Qualitätssicherung, die anders als die derzeit propagierten Verfahren des *data peer review* arbeiten. Transparenz und intersubjektive Nachvollziehbarkeit werden ebenso in den Forschungsprozess integriert wie die kontinuierliche Reflexion forschungsethischer Problematiken, etwa indem eine gemeinsame Interpretation von Daten in Arbeitsgruppen und Forschungswerkstätten stattfindet, in den Studien selbst die Entstehung der Daten reflektiert und der Interpretationszusammenhang möglichst transparent dargestellt wird. Die Monografie als verbreitetes Publikationsformat erlaubt neben der ausführlichen Entfaltung von Material die detaillierte Darlegung von Forschungsmethoden und Erkenntniswegen. Die Offenlegung von Interpretationen gegenüber beforschten AkteurInnen und Gruppen ist eine etablierte Möglichkeit, Ergebnisse diskursiv an das Forschungsfeld zurückzubinden.⁴

Mit Blick auf Standardisierung und Regulierung des Datenmanagements werden Passungsprobleme erkennbar, wenn etwa Standards der Datenpublikation aus anderen Disziplinen auf Forschungen aus der Europäischen Ethnologie übertragen werden, ohne deren methodische Besonderheiten, spezifische Forschungsstile und forschungsethische Implikationen ausreichend zu berücksichtigen. Deshalb sind gegenüber generischen Regeln differenzierende Verfahren der Langzeitarchivierung und Sekundärnutzung notwendig – nicht nur für die Europäische Ethnologie, sondern ähnlich auch für Nachbardisziplinen, in denen mit qualitativen Methoden geforscht wird.⁵ Entsprechend sollten Regelungen im Austausch mit Erfahrungen und Expertise etwa der Geschichtswissenschaften, der Ethnologie, der qualitativen Soziologie u.a., aber auch im Austausch mit Gedächtnisinstitutionen erarbeitet werden.

Insgesamt sind Überregulierung und Bürokratisierung der Prozesse vor allem für die ethnografische Feldforschung problematisch, denn Forschungsfragen und Zugänge zum Feld entwickeln sich häufig dynamisch. Nicht zuletzt weil relevante Akteur*innen im Vorhinein noch nicht bekannt sein können, können in der Konzeptionsphase nur begrenzt bzw. nicht abschließend Angaben zu den zu erwartenden Daten (Datentypen, Umfang, Relevanz für eventuelle Nachnutzung etc.) gemacht werden. Datenmanagementpläne – derzeit zentrales Instrument von Datenmanagement – können entsprechend bei An-

tragstellung lediglich eine grobe Orientierung abbilden. Prozesse der Konkretisierung sollen so gestaltet sein, dass frühere Entscheidungen durch nachfolgende differenziert und überprüft, aber auch revidiert werden können.

2. Rechtliche Fragen und Forschungsethik

Im Rahmen vor allem ethnografischer Forschungen entstehen zwar nicht immer, aber regelmäßig sensible, datenschutzrelevante Daten. Die Einwilligung der Beforschten zur Speicherung und Archivierung der Daten ist in diesen Fällen unabdingbar. Die Herstellung von Einverständnis kann aber je nach Forschungssituation auch ohne standardisierte schriftliche Vereinbarung – wie sie als dokumentierter *informed consent* konzipiert ist – erfolgen, sondern vielmehr prozessual, personengebunden und auf der Grundlage von bereits in der Ausbildung erlernten und in der Forschung permanent reflektierten forschungsethischen Standards. Letztere greifen auch dann, wenn Daten im datenschutzrechtlichen Sinn unbedenklich sind. Forschende verbürgen sich für einen verantwortungsvollen Umgang mit den ihnen anvertrauten Informationen – insbesondere im Veröffentlichungsfall, aber auch hinsichtlich des Verbleibs und der Verwendung der Daten nach dem Abschluss einer Forschung. Die Einhaltung gegebener Zusagen, Informationen vertraulich zu behandeln, ist in jedem Fall forschungsethisch geboten, ebenso wie die Vermeidung etwaiger Nachteile für beforschten Personen und Gruppen. Auch die Forderung, die Erhebung von Daten auf das Notwendige zu minimieren, wird im Rahmen des oben geschilderten Konzepts ethnografischen Arbeitens als Richtschnur gehandhabt. Die Flexibilität, den Prozess der Einwilligung angepasst an Bedingungen und Möglichkeiten des jeweiligen Forschungsfeldes zu gestalten, muss bewahrt werden, um Forschungsbeziehungen bzw. den Zugang zu Forschungsfeldern nicht zu beeinträchtigen und noch weiter zu erschweren. Das gilt – sofern die konkreten Daten das erforderlich machen – genauso für die Einholung von Einverständnis für die dauerhafte Archivierung und die Nachnutzung von Daten durch unbekannte Dritte. Auch hier sind ausreichend flexible Verfahren notwendig, vor allem hinsichtlich der Form und des Zeitpunkts der Zustimmung.

Verfahren der Anonymisierung bzw. Pseudonymisierung in Publikationen sind – sofern Forschungspartner*innen Anonymität wünschen – auch bisher ein zentrales Instrument des Persönlichkeitsschutzes und der Umsetzung forschungsethischer Anforderungen. Wege und Weisen der Anonymisierung von Daten in Forschungsprozessen sind jedoch vielfältiger und facettenreicher als die bisher im Kontext von Datenmanagement überwiegend diskutierten Verfahren. Das Spannungsverhältnis zwischen Anonymisierung, dem Erhalt der Interpretierbarkeit, die der Kontextualisierung von Daten bedarf, und damit des Werts der Daten für mögliche Nachnutzungen muss – auch angesichts des erheblichen Aufwands – sorgfältig und im Einzelfall reflektiert werden.

Insgesamt ist der Schutz der Interessen aller an einer Forschung beteiligten Personen bzw. Gruppen ein wichtiger Grundsatz, der Vorrang vor Verwertungsansprüchen hat. 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.

Datenarchivierung und Sekundärnutzung berühren darüber hinaus Fragen von Eigentums-, Urheberrechts- und Nutzungsrecht, die juristisch derzeit noch nicht geklärt sind. Auch innerfachlich gibt es unterschiedliche Auffassungen, inwiefern z. B. Daten aus ethnografischen Forschungen als von For-

schenden und Beforschten ko-produziert verstanden werden. Beobachtbar ist entsprechend ein Spektrum von voller Eigentümerschaft über die eigenen Aussagen beforschter Personen(gruppen) (im Sinne von geistigem Eigentum oder Kulturerbe) über ko-produzierte 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). Daten bleiben entsprechend im Besitz der Forschenden, können u. U. aber auch gemeinsam mit untersuchen Gruppen/ Personen gehalten werden, sofern das vereinbart wurde. Dies gilt insbesondere unter den Bedingungen akademischer Standortmobilität. Im Falle der Datenarchivierung in einem (Fach-)Repository muss auch diesem Umstand ausreichend Rechnung getragen werden.

Für die Bestimmung von Nutzungs- und Verwertungsrechten für alle Beteiligten müssen hier sowohl kurz- als auch längerfristig wirksame Mechanismen gefunden werden.

3. Archivierung und Nachnutzung. Szenarien und Bedingungen

Im Vordergrund des Interesses der Europäischen Ethnologie an der Datenarchivierung stehen nicht oder zumindest nicht in erster Linie Effizienz, Transparenz oder die Reproduzierbarkeit von Analysen, sondern das Vorhalten von Forschungsdaten für einen unbestimmten Zeitraum und wiederkehrende Nutzungen. Nachnutzungspotenziale könnten vor allem mit größerem zeitlichen Abstand zur Erhebung erkennbar werden, aber auch im Kontext von zunehmend kollaborativen Forschungen mehrerer Forscher*innen in einem Feld bzw. Themengebiet.

Weil Forschung in der Europäischen Ethnologie sowohl hinsichtlich ihrer Zugänge und den konkreten Bedingungen einzelner Forschungsfelder äußerst vielfältig ist, wird eine gleichmäßige, bedingungslose Verpflichtung zur Archivierung und Bereitstellung von Daten zur Nachnutzung seitens der *dgv* gleichwohl nicht befürwortet. Vielmehr muss die Freiwilligkeit der Datenüberlassung grundlegendes Prinzip sein (jenseits der geforderten Aufbewahrungspflicht gegenüber Drittmittelgebern). Forschende sollen unterstützt werden, die mit der Archivierung verbundenen Chancen zu nutzen, aber auch das Recht haben, die Nachnutzung der Daten einzuschränken: hinsichtlich des Umfangs der Daten, der Metadaten, der Codelisten, der Schutzfristen, des Personenkreises, der Zugang zu den Daten erhalten kann, der Art und Weise, wie die Daten zur Verfügung gestellt werden, sowie ggf. der Entwicklung von Regularien für die Datenlöschung.

Die Bedürfnisse der Primärforschung haben unbedingten Vorrang vor den Bedürfnissen eventueller Sekundärnutzungen. Daten sollen nicht mit Blick auf mögliche spätere Verwendung mit dazu erforderlichen Methoden erstellt werden. Es soll aber geprüft werden, welche Daten über Nachnutzungspotenzial verfügen und sich für eine Langzeitarchivierung eignen. Kriterien der Auswahl geeigneter Daten müssen im weiteren Prozess fachlicher Diskussion entwickelt werden. Sorgfältig abgewogen werden müssen zudem mögliche Nachteile, die insbesondere Forschungsteilnehmenden entstehen können. Die Abwägung ist überdies mit Blick auf verfügbare finanzielle wie personelle Ressourcen vorzunehmen.

Die Nachnutzung von Daten vor allem aus ethnografischen Forschungen soll in der Regel nur zu Forschungszwecken möglich sein, besonders in Hinblick auf möglichen *dual use* im weiteren Sinn sind eine sorgfältige Bewertung sowie ggf. Schutzmechanismen erforderlich. Eine Ausnahme können beforschte Gruppen bzw. Personen sein, sofern sie Zugang zu den mit ihnen gemeinsam hergestellten Materialien bekommen sollen. Hierüber sind individuelle Vereinbarungen zwischen Forschenden und Datenarchiven notwendig.

Ob und welche Materialien zur Nachnutzung freigegeben werden können, muss entsprechend einzel-fallbezogen entschieden werden. Eine regelrechte Veröffentlichung ethnografischer Daten wird in der überwiegenden Zahl der Fälle nicht möglich sein. Es muss möglich sein, Daten zu löschen, etwa wenn sich die Bedingungen, unter denen Daten an ein Repositorium übergeben wurden, ändern (wenn z. B. Einverständniserklärungen zurückgezogen werden).

Die Entscheidung über die Datenüberlassung an ein Repositorium kann nicht allein von Gremien oder Gutachtenden getroffen werden, sondern in erster Linie von Forschenden selbst auf der Grundlage der in die Forschung eingebetteten forschungsethischen Prinzipien. Die Entscheidung darüber wird im Verlauf des Forschungsprozesses getroffen und darf nicht förderrelevant sein.⁶

Insgesamt müssen Nachnutzungspotenziale und Aufwand im Einzelfall abgewogen werden. Der entstehende Mehraufwand muss zusätzlich zur bewilligten Forschungsförderung mit entsprechenden Ressourcen ausgestattet werden.

4. Abschließend

Die *dgv* unterstützt die Etablierung von Strukturen zur Langzeitarchivierung von Forschungsdaten sowie deren Bereitstellung für sekundäre Nutzungen, sofern beides freiwillig bleibt und methodische Vielfalt und Besonderheiten ethnografischer Forschungsprozesse angemessene Berücksichtigung finden. Die *dgv* setzt sich entsprechend dafür ein, dass der Heterogenität europäisch-ethnologischer Forschungspraxis bei der Etablierung disziplinübergreifender Regelungen und beim Aufbau von Datenarchiven und Repositorien ausreichend Rechnung getragen wird. Sie unterstützt Aktivitäten und Bemühungen, die auf den Aufbau von Kompetenzen und von notwendiger Sensibilität zielen, sowohl bei den Forschenden selbst als auch bei Institutionen, die künftig Daten aus ethnografischen Forschungen archivieren und bereitstellen können.

Weil die Veränderungen, die im Zusammenhang mit der Langzeitarchivierung und Nachnutzung von Daten im Rahmen eines professionellen Forschungsdatenmanagements angestoßen werden, nicht nur Arbeitsabläufe, rechtliche und technische Fragen betreffen, sondern sich im Hinblick auf epistemologische Vorannahmen, Forschungsethik und Methodologie sowie konkrete Forschungsdesigns in ihrer zeitlichen und praktischen Umsetzung auf den Forschungsprozess insgesamt auswirken, erfordert ein adäquates Datenmanagement perspektivisch auch spezifische Beratung und Schulung sowie neue curriculare Inhalte in der Ausbildung von Studierenden und Promovierenden. Die *dgv* fordert Hochschulen, Bildungspolitik und Wissenschaftsorganisationen dazu auf, die notwendigen Mittel dafür sowie für die Aufbereitung und Sicherung von Forschungsdaten zur Verfügung zu stellen und geeignete Infrastrukturen aufzubauen.

Ergänzend: Allgemeine Anforderungen an künftige Datenarchive

Derzeit existieren im deutschsprachigen Raum bestenfalls Ansätze zur Etablierung von Datenarchiven, die den forschungsethischen wie rechtlichen Anforderungen qualitativer Forschung in der gesamten Breite gerecht werden können. Benötigt werden spezialisierte, in hohem Maße professionalisierte Datenarchive (Fachrepositorien) in öffentlicher Hand, die langfristig abgesichert arbeiten und wissenschaftlich unabhängig geführt werden müssen.

Datengebende sollten über Nachnutzungen und ggf. über Publikationen informiert werden. Dass eine Zustimmung der Datengebenden zur Nutzung erforderlich ist, kann im Einzelfall vereinbart werden.

Zugangsrechte müssen je nach Schutzbedarf der Daten differenziert und dauerhaft kontrolliert vergeben werden, einschließlich der Einrichtung speziell geschützter Räume im Archiv selbst, in denen Daten mit hohem Schutzbedarf eingesehen werden können – so wie das in einzelnen Datenzentren bereits üblich ist (*safe room*). Rechte und Pflichten von Nutzenden müssen – ähnlich wie das in vielen ‚analogen‘ Archiven üblich ist – vertraglich vereinbart werden. Weil technische Schutzmaßnahmen nicht genügen, wird ausreichend und geschultes Personal benötigt.

Benötigt werden Exit-Strategien: Datenarchive müssen die Löschung von Daten garantieren, wenn sich die Bedingungen ändern, unter denen Daten übergeben wurden. Sie müssen erklären, was mit den Daten geschieht, falls die Archive selbst aufgelöst werden sollten.

Stand: 19.11.2018

¹ Das Thema Datenmanagement und Entwürfe dieses Papiers sind auf dem Weg unterschiedlicher Beteiligungsformen und Veranstaltungen der *dgv* diskutiert und erarbeitet worden. Das vorliegende Papier wurde am 23. November 2018 von Vorstand und Hauptausschuss der *dgv* verabschiedet. Autorinnen des Papiers sind vor allem Beate Binder, Sabine Imeri, Gertraud Koch und Gisela Welz. Es geht zurück auf einen Diskussionszusammenhang, an dem auch Ulrike Klöppel, Elisabeth Huber und Thomas Stodulka beteiligt waren. Eingeflossen sind überdies Überlegungen und Kommentare von weiteren Fachkolleg*innen, insbesondere von Martina Klausner.

² http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf

³ Im Weiteren wird verkürzend von Europäischer Ethnologie gesprochen. Der „Fachzusammenhang“ umfasst universitäre wie außeruniversitäre Institute in der Fachtradition von Volkskunde, Empirischer Kulturwissenschaft, Europäischer Ethnologie, Kulturanthropologie.

⁴ Teilweise werden inzwischen auch Forschungsstrategien mit *open data* Ansätzen exploriert, in denen die Akteur*innen in den Forschungsfeldern auf neue Weise in die Datengenerierung eingebunden werden.

⁵ Die Notwendigkeit der Differenzierung erkennen auch Wissenschaftspolitik beratende Gremien wie der Rat für Sozial- und Wirtschaftsdaten an. Vgl. z.B. Archivierung und Sekundärnutzung von Daten der qualitativen Sozialforschung. Eine Stellungnahme des RatSWD. (2015) http://www.ratswd.de/dl/RatSWD_Output1_Qualidaten.pdf

⁶ Im Sinne einer „Ethik der Nachnutzung“ ist die Diskussion von Richtlinien/Standards für den Umgang mit Daten erforderlich, die von anderen Wissenschaftler*innen erhoben worden sind.