

Stellungnahme des RatSWD zur Archivierung und Sekundärnutzung von Daten der qualitativen Sozialforschung

Zusammenfassung

Der RatSWD will optimale Bedingungen für die Genese von und den Zugang der Wissenschaft zu Forschungsdaten herstellen. Er befürwortet grundsätzlich die Archivierung und die Bereitstellung qualitativer Daten für Sekundäranalysen. Der RatSWD erkennt an, dass die Frage der Möglichkeit, der Angemessenheit und des wissenschaftlichen Nutzens von Sekundäranalysen nach Materialart und Forschungsmethoden differenziert zu beantworten ist. Er spricht sich dafür aus, auch im Bereich der qualitativen Sozialforschung grundsätzlich eine Kultur der Datenbereitstellung zu fördern. Hier sind Forschende typischerweise aktiv in die Produktion von Datenmaterial eingebunden, so dass die Frage der Datenweitergabe die Frage des geistigen Eigentums berührt. Es gilt, Archivierungsverfahren bereit zu stellen, die versprechen, den Wert des Datenmaterials zu erhalten, ohne den Datenschutz zu gefährden. Regelungen zur Archivierung wie zur Sekundärnutzung dürfen den Feldzugang für Primärforschende nicht einschränken. Der RatSWD befürwortet die Archivierung qualitativer Daten und ihre Bereitstellung für Sekundäranalysen soweit dies ohne nachvollziehbare Gefährdung der primären Forschungsziele möglich ist. Bei Drittmittelprojekten soll die Vorlage eines Datenmanagementplans bei der Beantragung obligatorisch gemacht werden. Die Entscheidung über die Eignung von Daten für eine Sekundärnutzung darf keinen Einfluss auf die Genehmigung beantragter Projekte haben.

Ziele des RatSWD

Der RatSWD ist dem Ziel verpflichtet, optimale Bedingungen für die Schaffung von und den Zugang der Wissenschaft zu Forschungsdaten herzustellen. Daher unterstützt der RatSWD Bemühungen um eine nachhaltige Sicherung und Bereitstellung von qualitätsgesicherten Forschungsdaten - insbesondere aus öffentlicher Förderung - für die Wissenschaft. Solche Daten sind in vielen Forschungsfeldern eine Voraussetzung für wettbewerbsfähige und innovative Forschung. Hierbei spielen die Entwicklung und Einhaltung methodenangemessener Standards zur Sicherstellung einer hohen Datenqualität – auch mit Blick auf die Wahrung geistigen Eigentums, die Innovationsfähigkeit des Forschungssystems, datenschutzrechtliche und forschungsethische Anforderungen – ebenso wie die praktische Weiterentwicklung und Verbesserung entsprechender Forschungsinfrastrukturen eine maßgebliche Rolle.

Entwicklung der Qualitätssicherung in der empirischen Sozial- und Wirtschaftsforschung

Der RatSWD nimmt insbesondere im Bereich der quantitativen Analysen (bspw. in der empirischen Wirtschaftsforschung und in vielen anderen empirischen Disziplinen) Entwicklungen hin zu einer Replikationskultur zur Kenntnis. In diesem Rahmen wird die Bereitstellung von Analysedaten oder wenigstens Programmcodes, die auf zugangsgeschützte Daten nach Klärung von Zugangsrechten angewendet werden können, zum wissenschaftlichen Standard. Neben den Kostensparnissen, die durch die Bereitstellung vorhandener Daten für weitere Forschende realisiert werden können, und die damit verbundene Erleichterung von Anschluss- und Parallelforschung, erleichtert dieser Kulturwandel die Prävention und das Aufdecken von Fehlern und führt zu einem bewussteren Umgang mit und der sorgfältigeren Analyse von Daten. Zur besseren Erreichbarkeit dieser Ziele wurden in Deutschland für die quantitativen Sozial-, Verhaltens- und Wirtschaftswissenschaften bislang fast 30 Forschungsdaten- und Servicezentren eingerichtet. Die Zahl von vergleichbaren Einrichtungen wächst auch in anderen Ländern.

Ziel: Archivierung und Bereitstellung von qualitativen Daten für Sekundärnutzung

Der RatSWD befürwortet grundsätzlich die Archivierung und die Bereitstellung qualitativer Daten für Sekundäranalysen und unterstützt die aktuell laufenden Initiativen und Forschungsprojekte, die sich mit der potenziellen Sekundärnutzung qualitativer Daten beschäftigen. Wissenschaftliche Aussagen sollten auf transparente und für Dritte nachvollziehbare Weise zustande kommen. In der qualitativen Forschung wird dies praktiziert durch kollektive Formen der Ergebnisproduktion in Interpretationsgruppen und kooperativen Settings sowie durch die Offenlegung von Qualitätssicherungsmaßnahmen im Forschungsprozess, bspw. in Berichten und Publikationen. Zudem gibt es, wie in der quantitativen Forschung, die Möglichkeit der Reproduzierbarkeit durch erneute Forschung; klassische Re-Analysen, wie sie sich in der quantitativen Forschung bewährt haben, sind in der qualitativ-interpretativen Forschung aus verschiedenen – insbesondere methodologischen – Gründen in der Regel nicht möglich. Transparenz über die Her- und Ableitung wissenschaftlicher Aussagen ist in beiden wissenschaftlichen Traditionen ein unverzichtbares konstituierendes Element wissenschaftlicher Forschung. Daher unterstützt der RatSWD die Bemühungen der Fachgesellschaften und der Forschungsförderer um einen differenzierten eigentumsrechtlich angemessenen, ethischen und datenschutzrechtlichen Umgang mit der Vielfalt an Datentypen. Dies entspricht den internationalen Standards zur Archivierung und Sekundärnutzung von Daten, die Forschende auffordern zu prüfen, ob und mit welchen speziellen Arrangements qualitative Daten für eine Sekundärnutzung bereitgestellt werden können.¹

In Übereinstimmung mit der gemeinsamen Stellungnahme der Sektionen „Biographieforschung“ und „Methoden der qualitativen Sozialforschung“ der Deutschen Gesellschaft für Soziologie (DGS)² erkennt der RatSWD an, dass die Frage der Möglichkeit, der Angemessenheit und des wissenschaftlichen Nutzens von Sekundäranalysen nach Materialart und Forschungsmethoden differenziert zu beantworten ist. Der RatSWD ist sich der Tatsache bewusst, dass Datenmaterial in der qualitativen Sozialforschung in Prozessen generiert wird, die sich von denen der quantitativen Sozialforschung deutlich unterscheiden. Gleichwohl spricht er sich dafür aus, auch im Bereich der qualitativen Sozialforschung grundsätzlich

¹ Eine allgemeine Diskussion sowie zahlreiche Verweise finden sich bei Cliggett (2013, <http://www.nova.edu/ssss/QR/QR18/cliggett1.pdf>). Die amerikanische National Science Foundation beschreibt Vorgaben zu "Data Archiving Policy" für qualitative Daten (www.nsf.gov/sbe/ses/common/archive.jsp) und auch das britische ESRC verweist in seiner Research Data Policy auf qualitative Daten (http://www.esrc.ac.uk/_images/research-data-policy_tcm8-34123.pdf).

² http://www.soziologie.de/fileadmin/user_upload/Sektion_Biographieforschung/Resolution_Datenarchivierung_Final-1.pdf

eine Kultur der Datenbereitstellung zu fördern. Anträge auf öffentlich finanzierte Forschungsförderung sollten Aussagen zur Möglichkeit der Datenbereitstellung für eine Sekundärnutzung treffen. Forschungsförderer sollten bei der Bewertung der Förderanträge prüfen bzw. von den Gutachtenden beurteilen lassen, ob die zu sammelnden qualitativen Daten potenziell für eine Sekundärnutzung geeignet und nützlich wären und ob die Antragstellenden gegebenenfalls angemessene Maßnahmen der Archivierung und Bereitstellung (im Rahmen des Datenschutzes) vorsehen. Da die Förderung wissenschaftlicher Forschung vorrangig eine Frage fachlicher Qualität und inhaltlicher Relevanz ist, soll die Förderung nicht von der Eignung der Daten für Sekundärnutzungen abhängig gemacht werden.

Auch im Bereich der qualitativen Forschung in den Sozial- und Wirtschaftswissenschaften soll den Forschenden eine Infrastruktur verfügbar gemacht werden, die eine sichere Archivierung der produzierten Forschungsdaten erlaubt, etwa in Form spezialisierter Forschungsdatenzentren. „Sicher“ meint hier – wie im Bereich der quantitativen Forschung – eine langfristig verlässliche technische Qualität, aber auch Schutz vor missbräuchlichen Zugriffen durch Dritte.

Zu klären ist im Zusammenhang mit Archivierung und Sekundärnutzungen auch der Umgang mit schutzwürdigen Eigentumsrechten an den Informationen, die in qualitativen Studien enthalten sind bzw. die im Zuge einer qualitativen Studie von Primärforschenden erzeugt werden. Die Klärung und der Schutz von Eigentumsrechten ist hierbei kein Spezialproblem der qualitativen Forschung, sondern sie ist ein übergreifendes Problem in der gesamten Scientific Community.

Spezifika qualitativer Forschung

Innerhalb der qualitativen Sozialforschung werden – ebenso wie inzwischen auch in der quantitativen Sozial- und Wirtschaftsforschung – verschiedene, gleichermaßen wissenschaftlich legitimierte Forschungsstile und -methoden praktiziert, die unterschiedliche Typen von Datenmaterial hervorbringen.

Qualitative Sozialforschung erforscht häufig sensible Bereiche und Prozesse der Gesellschaft, zu denen ein produktiver unmittelbarer Zugang nur herzustellen ist, wenn ein besonderes Vertrauensverhältnis zwischen Forschenden und Informanten und Informantinnen etabliert werden kann (z.B. Devianzforschung oder Migrationsforschung). Die Ankündigung möglicher Sekundärnutzungen vermag hier nicht nur den Zugang, sondern auch den Rapport mit dem Feld empfindlich zu stören.³

³ Zu den datenschutzrechtlichen Anforderungen bei der Generierung und Archivierung qualitativer Interviewdaten vgl. Liebig et al. 2014 (Working Paper 238/2014 des RatSWD unter http://www.ratswd.de/dl/RatSWD_WP_238.pdf).

Datenmaterial der qualitativen Forschung wird für die jeweils spezifischen Untersuchungsfragen generiert, die in starkem Maße die Art und die Auswahl des Datenmaterials prägen. Sekundärnutzungen für andere als die Ausgangsfragestellungen werfen daher gerade bei stark kontextabhängigen Daten (z.B. in der Biografieforschung oder der Ethnographie) methodologische Probleme auf, die erheblichen Anforderungen an die Archivierung und an eine Freigabe der Daten für Zwecke anderer Forschender stellen. Kontexte, die eine Bewertung des Datenmaterials erlauben, müssen dazu dokumentiert werden.

Forschende sind in unterschiedlichem Maße, in der qualitativen Forschung jedoch typischerweise sehr ausgeprägt, aktiv in die Produktion von Datenmaterial eingebunden. Insbesondere in ethnographischen Verfahren werden sie zu Autoren ihres Materials, das umgekehrt immer bereits Elemente der analytischen Eigenleistung der Forschenden enthält. Die Frage der Weitergabe berührt in diesen Fällen also in besonderem Maße die Frage des geistigen Eigentums.

Aus der Forschungslogik qualitativer Forschung folgen Konsequenzen für die Art und den Status der Daten und damit auch für Fragen der Archivierung und Sekundärnutzung.

Anforderungen an die Archivierung und Sekundärnutzung qualitativer Daten

Besondere Anforderungen ergeben sich zunächst mit Blick auf Möglichkeiten, Erfordernisse und Aufwand einer dem Material angemessenen Anonymisierung und Beschreibung des Kontextes der Daten (Meta-Daten). Bedürfen schon die kontextreichen Daten von Leitfadeninterviews einer umfänglichen Maskierung, um sie ohne Gefährdung des Schutzes der Informanten und Informantinnen Dritten zugänglich machen zu können, so nimmt dieser Aufwand bei audio-visuellen Daten und bei umfassenden Feldprotokollen ethnographischer Forschung erheblich zu. Zugleich schränken die erforderlichen Anonymisierungsschritte (z.B. verpixelte Gesichter bei videographischen Interaktionsanalysen) den heuristischen Wert des Materials erheblich ein. Es gilt zunächst, für die zur Sicherung der guten wissenschaftlichen Praxis erforderlichen Aufbewahrungsfristen Archivierungsverfahren bereit zu stellen, die versprechen, den Wert des Datenmaterials für die Forschenden zu erhalten, ohne den Datenschutz zu gefährden.

Sekundärnutzungen qualitativer Daten außerhalb des ursprünglichen Datenkontextes stellen darüber hinausgehende Anforderungen und werfen über die Archivierungsfragen hinaus zusätzliche Fragen auf. Dies betrifft mindestens die folgenden Aspekte:

Regelungen zur Archivierung wie zur Sekundärnutzung dürfen die Möglichkeiten zum Feldzugang und damit zur Gewinnung relevanter Daten nicht negativ beeinflussen. Die Produktion von Daten hat Vorrang vor deren Sekundärnutzung. Während eine potentielle Sekundärnutzung in einigen Forschungsfeldern (z.B. Devianzforschung, fallbezogene Studien in der Medizinsoziologie, Managementstudien, Biographieforschung) den Feldzugang und damit die Datenproduktion offenkundig unterbinden kann, besteht derzeit abgesehen von den Erfahrungen der Forschenden wenig empirisch gesichertes Wissen, wie sich dies in anderen Forschungsfeldern auf die Teilnahme- und Auskunftsbereitschaft auswirkt. Der RatSWD sieht hier einen besonderen Forschungsbedarf.

Sollen Sekundärnutzungen gewinnbringend möglich sein, so bedarf es nicht allein der Bereitstellung des Datensatzes; zusätzlich ist eine umfangreiche und detaillierte Dokumentation des gesamten Materialbestandes, der Umstände seiner Genese sowie der Art und Weise, wie von dem Material im Einzelnen analytisch Gebrauch gemacht wurde, erforderlich. Dies geht über den bislang zu leistenden Aufwand im Rahmen guter wissenschaftlicher Praxis in der qualitativen Forschung weit hinaus und stellt, da dies nicht von Dritten geleistet werden kann, eine erhebliche zusätzliche Belastung der Primärforschenden dar.

Die Überlassung von Daten aus der qualitativen Forschung berührt eine grundlegende Frage des Urheberrechts an empirischen Daten. Urheberrechte lassen sich generell nicht abtreten. Vielmehr geht es immer um die Übertragung von Nutzungsrechten und die Kompensation des Eigentümers für diese Abtretung. Diese Kompensation erfolgt in der Wissenschaft normalerweise durch Zitation des Eigentümers.⁴ Diese Art der Kompensation in Bezug auf Daten ist in den Wissenschaftsdisziplinen noch nicht hinreichend geklärt und weltweit in der Diskussion. Dabei geht es auch um die zentrale Frage, wer die Entscheidung bezüglich der Nutzungsrechte treffen darf (Primärforschende, Mitarbeitende in Forschungsteams, Einbezug von drittmittelgebenden Stellen) und ab wann konkurrie-

⁴ Die Vergabe von DOIs ist der etablierte Standard für digitale Objekte (Forschungsdaten, elektronische Fassungen von Artikeln etc.), um sie mit einem persistenten Identifikator zu versehen. Der Standard wird von der DOI Foundation getragen. Für unterschiedliche digitale Objekte gibt es unterschiedliche Registrierungsagenturen. Für Forschungsdaten ist das DataCite (unter Führung der Technischen Informationsbibliothek) zuständig. GESIS betreibt gemeinsam mit der Deutschen Zentralbibliothek für Wirtschaftswissenschaften die Registrierungsagentur für Sozial- und Wirtschaftsdaten, diese ist wiederum Mitglied im DataCite-Verbund. Die da|ra bietet einen DOI-Registrierungsservice und einen Metadatenverwaltungsdienst an. Datenerhaltende Einrichtungen (FDZs und DSZs) können mit der da|ra ein Service-Level-Agreement abschließen, um Daten zu registrieren (Registrierungsagenten). Alle in Deutschland bekannten und etablierten Infrastruktureinrichtungen für qualitative Daten (Qualiservice, DIPF, DSZ-BO u. a.) registrieren ihre Daten bereits nach dem DOI Standard. Die Empfehlungen von GESIS zum bibliographischen Zitieren von Forschungsdaten finden sich unter <http://www.gesis.org/unsere-angebote/daten-analysieren/datenservice/forschungsdaten-zitieren/> (siehe auch <http://www.da-ra.de/>).

renden Forschende ein Recht auf Sekundärnutzung zum Zwecke der Prüfung durch Re-Analysen und für eigene Forschungsfragen eingeräumt wird.

Empfehlungen

Der RatSWD befürwortet die Archivierung qualitativer Daten und ihre Bereitstellung für Sekundäranalysen soweit dies ohne nachvollziehbare Gefährdung der primären Forschungsziele möglich ist. In der öffentlich geförderten qualitativen Sozial- und Wirtschaftsforschung soll – wie in der quantitativen Forschung – grundsätzlich eine Kultur der Archivierung und Datenbereitstellung für Sekundäranalysen gefördert werden. In Forschungsanträgen sollen auch Angaben zum mittel- und langfristigen Umgang mit Forschungsdaten gemacht werden (Datenmanagementplan). Einschränkungen, die sich bereits aus Regelungen des Datenschutzes ergeben, bleiben davon unbenommen.

Die Entscheidung, ob und wie Forschungsdaten für Sekundäranalysen genutzt werden können – etwa zur Prüfung publizierter Ergebnisse, Nutzung für mehr oder weniger stark von der Primärnutzung abweichenden Forschungsfragen – sollte von Primärforschenden, Gutachtenden und Förderinstitution gemeinsam im Prozess der Projektentwicklung und -begutachtung nach Möglichkeit vor Beginn eines Forschungsvorhabens, spätestens aber bei Projektabschluss getroffen werden. Dabei werden typischerweise – wie dies zum Beispiel in den Lebenswissenschaften gängig ist – auch Fristen für die Bereitstellung der Daten definiert, die die berechtigten Interessen der Datenproduzenten an der Verwertung der Forschungsergebnisse für eigene Publikationen wahren. Die Entscheidung über die Eignung der Daten für eine Sekundärnutzung darf keinen Einfluss auf die Genehmigung beantragter Projekte haben.

Bei Drittmittelprojekten soll die Vorlage eines Datenmanagementplans bei der Beantragung obligatorisch gemacht werden; der Plan ist in den Peer-Review-Prozess einzubeziehen. Dieser Plan kann ggf. auch begründet vorsehen, dass erst im Projektverlauf und abhängig von den Erfordernissen des Forschungsprozesses konkret über Details des Datenzugangs entschieden wird. Auf diese Weise wird die Frage, ob aus forschungspraktischen, datenschutzrechtlichen oder ethischen Gründen Sekundäranalysen nicht möglich oder nicht ratsam sind, transparent innerhalb der geeigneten Scientific Community geklärt. Bei Drittmittelanträgen, bei denen eine Sekundärnutzung nicht möglich erscheint oder nicht ratsam ist, sollten die Drittmittelgeber und Gutachter ausdrücklich Anträge für Nachfolgeprojekte zulassen, die mit Hilfe neuer Datenerhebungen die Ergebnisse des Ursprungsprojektes prüfen.

Die Frage nach den Eigentumsrechten an Forschungsdaten bedarf einer grundsätzlichen Klärung. Insbesondere müssen die Scientific Communities sich darüber verständigen und geeignete Regeln implementieren, wie die Sekundärnutzung von Forschungsdaten dem Datenproduzenten oder der Datenproduzentin entgolten werden. Dabei geht es in der Regel nicht um eine monetäre Kompensation, sondern um die angemessene Zitation des Primärdatenproduzenten.

Der RatSWD drückt darüber hinaus die Erwartung aus, dass für alle Forschungsdaten der für sekundäranalytische Nutzungsmöglichkeiten erforderliche Mehraufwand von den Projektträgern (insb. akademischen Drittmittelgebern) angemessen vergütet wird, sodass dadurch weder die Finanzierung der eigentlichen Forschungsaufgaben eingeschränkt wird, noch die Bewilligungsquoten insgesamt abgesenkt werden.