

Auszug aus dem Buch:

Uwe Jensen  
Sebastian Netscher  
Katrin Weller (Hrsg.)

# Forschungsdatenmanagement sozialwissenschaftlicher Umfragedaten

Grundlagen und praktische Lösungen  
für den Umgang mit  
quantitativen Forschungsdaten

Verlag Barbara Budrich  
Opladen • Berlin • Toronto 2019

Bibliografische Information der Deutschen Nationalbibliothek  
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie;  
detaillierte bibliografische Daten sind im Internet über  
<http://dnb.d-nb.de> abrufbar.

© 2019 Dieses Werk ist beim Verlag Barbara Budrich erschienen und steht unter der Creative Commons Lizenz Attribution-ShareAlike 4.0 International (CC BY-SA 4.0):

<https://creativecommons.org/licenses/by-sa/4.0/>.

Diese Lizenz erlaubt die Verbreitung, Speicherung, Vervielfältigung und Bearbeitung bei Verwendung der gleichen CC-BY-SA 4.0-Lizenz und unter Angabe der UrheberInnen, Rechte, Änderungen und verwendeten Lizenz.



Dieses Buch steht im Open-Access-Bereich der Verlagsseite zum kostenlosen Download bereit (<https://doi.org/10.3224/84742233>).

Eine kostenpflichtige Druckversion (Print on Demand) kann über den Verlag bezogen werden. Die Seitenzahlen in der Druck- und Onlineversion sind identisch.

ISBN 978-3-8474-2233-4 (Paperback)

eISBN 978-3-8474-1260-1 (eBook)

DOI 10.3224/84742233

Umschlaggestaltung: Bettina Lehfeldt, Kleinmachnow – [www.lehfeldtgraphic.de](http://www.lehfeldtgraphic.de)

Lektorat: Nadine Jenke, Potsdam

Satz: Anja Borkam, Jena – [kontakt@lektorat-borkam.de](mailto:kontakt@lektorat-borkam.de)

Titelbildnachweis: Foto: Florian Losch

Druck: paper & tinta, Warschau

Printed in Europe

## 7. Data Sharing: Von der Sicherung zur langfristigen Nutzung der Forschungsdaten

*Reiner Mauer und Jonas Recker*

Transparente Forschungsprozesse und reproduzierbare Ergebnisse sind Qualitätsmerkmale wissenschaftlichen Arbeitens. Die dauerhafte Verfügbarkeit von Forschungsdaten – und damit auch deren Archivierung als notwendige Voraussetzung – leistet hierbei einen wesentlichen Beitrag: Sie macht empirische Forschung nicht nur nachvollziehbar, sondern auch anschlussfähig. Darüber hinaus ist ein möglichst offener Zugang zu qualitätsgesicherten und gut dokumentierten Forschungsdaten eine wichtige Basis für den wissenschaftlichen Fortschritt, indem Möglichkeiten für neue und innovative Fragestellungen auch über Disziplinengrenzen hinweg geschaffen werden.

Dieses Bewusstsein für den Wert von Forschungsdaten ist im gesamten Wissenschaftssystem deutlich gestiegen, mit teilweise sehr konkreten Auswirkungen: Es schlägt sich zum einen in gesteigerten förderpolitischen Maßnahmen zum Aufbau datenbezogener Infrastrukturen und Services nieder und in der Folge in einer Zunahme entsprechender Angebote. So sind beispielsweise in den letzten Jahren an vielen Hochschulen, aber auch an außeruniversitären Einrichtungen Repositorien und Services für Forschungsdaten aufgebaut worden, um den Bedarf der Forschenden an professionellen Lösungen für das Management, die Sicherung und die Veröffentlichung von Forschungsdaten zu decken. Darüber hinaus treten neue, teils kommerzielle Anbieter wie Figshare oder Mendeley Data in Erscheinung.

Auch sehen Forschende sich zunehmend mit gestiegenen Erwartungen hinsichtlich des Umgangs mit den von ihnen erzeugten Forschungsdaten konfrontiert. Das veränderte Bewusstsein für den Wert von Forschungsdaten drückt sich auf forschungs- und förderpolitischer Ebene mittlerweile nicht mehr nur in der Form von eher allgemeinen und unverbindlichen Appellen aus, sondern mündet zunehmend auch in konkreten Maßnahmen und Auflagen. So gehen Forschungsförderer etwa dazu über, zumindest Aussagen zum geplanten Umgang mit erzeugten Daten bereits bei der Antragsstellung einzufordern, und beziehen diese zunehmend in die Entscheidung über eine Förderung ein. Vereinzelt finden sich mittlerweile in Bescheiden deutscher Förderer Auflagen dazu, wie nach Projektende mit den erzeugten Forschungsdaten zu verfahren ist. Zum Beispiel fordert das Bundesministerium für Bildung und Forschung (BMBF) in bestimmten Programmen, dass „Standards des Forschungsdatenmanagements“ eingehalten und erhobene Forschungsdaten nach Projektende an ein geeignetes Forschungsdatenzentrum übergeben werden, „um im Sinne der guten wissenschaftlichen Praxis eine langfristige Datensicherung für Replikationen und gegebenenfalls Sekundärauswertungen zu ermöglichen“ (BMBF 2017: B2).

Auch wenn derartige Regelungen zurzeit in Deutschland eher noch die Ausnahme bilden, lässt sich eine Tendenz hin zu mehr Verbindlichkeit – wie sie beispielsweise in den USA und im Vereinigten Königreich schon deutlich weiter vorangeschritten ist – auch in Deutschland erkennen. Angesichts dieser Entwicklung stellt sich für viele Forschende die (idealerweise schon zu Projektbeginn zu beantwortende) Frage, wie sie mit ihren Forschungsdaten nach dem Ende des Projekts verfahren sollen. In diesem Zusammenhang gilt es, die geeignete Form der Sicherung, Archivierung oder Veröffentlichung der Daten zu bestimmen und sich ggf. für ein geeignetes Repository oder Datenzentrum zu entscheiden.

In diesem Kapitel soll ein Überblick über die verschiedenen Optionen gegeben werden, Forschungsdaten zu erhalten und zur Nachnutzung durch Dritte zur Verfügung zu stellen. Dazu werden zunächst zentrale Begriffe erläutert, die Forschenden immer wieder begegnen, wenn sie sich mit der Frage beschäftigen, was mit den Daten nach Ende des eigenen Forschungsprojekts geschehen soll (Abschnitt 7.1). Abschnitt 7.2 befasst sich darauf aufbauend mit Zugangswegen zu Forschungsdaten sowie mit Lizenzen, die die Nutzung der Daten regeln können. Anschließend soll in Abschnitt 7.3 ein Überblick über bestehende Möglichkeiten gegeben werden, Daten über die Projektlaufzeit hinaus zu sichern und zugänglich zu machen. Abschnitt 7.4 beschreibt Maßnahmen, die Datenzentren üblicherweise durchführen, wenn sie Forschungsdaten für eine langfristige Sicherung und Bereitstellung übernehmen. Das Kapitel schließt mit Hinweisen dazu, wie Forschende ein für ihre Bedarfe geeignetes Repositorium auswählen können (Abschnitt 7.5).

## 7.1 Zentrale Begriffe: Forschungsdaten sichern, archivieren und teilen

Das Thema Forschungsdatenmanagement ist eng mit der Frage verknüpft, was mit Forschungsdaten geschehen soll, wenn das Projekt endet, in dem sie erhoben und analysiert wurden. In diesem Zusammenhang begegnen Wissenschaftlerinnen und Wissenschaftlern einer Vielzahl von Begriffen. So ist die Rede davon, dass Daten aufbewahrt, gesichert, archiviert und geteilt werden sollen. In der Regel existieren aber keine einheitlichen Definitionen der Begriffe und ihre Bedeutung ist stark vom jeweiligen Verwendungskontext abhängig. Dies erschwert das Verständnis sowohl in Bezug auf die Anforderungen, denen Forschende genügen sollen, aber auch im Hinblick auf das Leistungsspektrum verschiedener Datenzentren. Wie unterscheidet sich z.B. eine *Sicherung* von einer *Archivierung* der Daten? Und was bedeutet es, Daten mit Dritten zu teilen (*Data Sharing*)?

Die nachfolgende Klärung häufig verwendeter Terminologie soll vor diesem Hintergrund dazu beitragen, ein gemeinsames Verständnis für die in diesem Kapitel verwendeten Begrifflichkeiten zu entwickeln. Insbesondere sollen die Erläuterungen Forschenden aber dabei helfen, forschungsdatenbezogene Empfehlungen und Forderungen (vgl. z.B. DFG 2015; Allianz der deutschen Wissenschaftsorganisationen 2010) besser zu verstehen und Serviceangebote am Markt in ihrem Leistungsspektrum einschätzen zu können.

### 7.1.1 Aufbewahrung und Sicherung

Mit der Forderung, Forschungsdaten *aufzubewahren* bzw. zu *sichern* (vgl. DFG 2013: Empfehlung 7), gehen die geringsten Anforderungen an den Umgang mit den Objekten (in unserem Fall digitale Forschungsdaten und dazugehörige Begleitmaterialien, wie z.B. die Datendokumentation) einher. Im Vordergrund der Maßnahmen steht typischerweise der Erhalt dieser Objekte in definierten (in der Regel unveränderten) Zuständen. Dies beinhaltet vor allem den Schutz vor Verlust – sei es durch Zerstörung oder Löschung, Beschädigung und damit einhergehender Unlesbarkeit oder durch Unauffindbarkeit – sowie den Schutz vor unberechtigtem Zugriff. Die Sicherung digitaler Objekte wird insbesondere durch die systematische Speicherung und weitere organisatorische und technische Maßnahmen realisiert.

### 7.1.2 (Digitale) Archivierung und Kuratierung

Im Vergleich zur Sicherung ist mit dem Begriff der *Archivierung* nicht nur ein systematisches Vorgehen bei der Speicherung bzw. im analogen Umfeld der Aufbewahrung von Objekten verbunden, sondern es kommen – der Tradition von Registraturen und Archiven folgend – weitere zentrale Funktionen hinzu, wie etwa Auswahl und Bewertung, Erschließung, Herstellung und Erhalt der Nutzbarkeit sowie die Organisation des Zugriffs auf das Archivgut. Typischerweise zielt Archivierung im Kontext von Bibliotheken, Archiven oder Museen darauf ab, Objekte zeitlich unbegrenzt aufzubewahren, benutzbar zu machen und zu erhalten. Die zentralen Konzepte und Verfahrensweisen der Archivierung wurden zwar im analogen Umfeld entwickelt, finden mittlerweile aber auch im Bereich der *digitalen Archivierung* (auch: digitale Langzeitarchivierung) Anwendung. Die Erhaltung von Information in digitaler Form gestaltet sich dabei wesentlich schwieriger als der Erhalt analoger Objekte, wie z.B. Papier oder Mikrofilm. Dies ergibt sich vor allem daraus, dass digitale Informationen abhängig von komplexen, technischen Umgebungen sind. Digitale Informationen sind mehr oder weniger eng mit Hardware, Speichermedien, Betriebssystemen und Anwendungssoftware verbunden. Diese unterliegen einem ständigen technologischen Wandel und so ist es in erster Linie der technologische Fortschritt, der digitale Informationen kontinuierlich bedroht. Eine gesicherte Aufbewahrung digitaler Informationen im Sinne von Speicherung reicht demnach nicht aus, um deren Lesbarkeit und schon gar nicht deren Interpretierbarkeit langfristig sicherzustellen. *Langfristig* wird dabei im zentralen Rahmenstandard für die digitale Langzeitarchivierung (ISO 14721: Open Archival Information System [OAIS]: Reference Model) verstanden als „lange genug [...], um sich mit den Auswirkungen des Technologiewandels inklusive der Unterstützung von neuen Datenträgern und Datenformaten sowie einer sich verändernden vorgesehenen Zielgruppe“ zu befassen (nestor 2013: 13).

Digitale Langzeitarchivierung verfolgt daher im Wesentlichen zwei sich ergänzende Ziele: (1) die Substanzerhaltung der Dateninhalte, aus denen digitale Objekte physikalisch bestehen (*bitstream preservation*) sowie (2) den Erhalt der dauerhaften Benutzbarkeit und (inhaltlichen) Interpretierbarkeit. Letzteres zielt darauf ab, die in den Daten enthaltenen Informationen sowie das für deren Interpretation notwendige Wissen, das nicht in den Daten selbst enthalten ist, zu sichern. So wird inhaltliche Interpretierbarkeit gewährleistet, indem alle zum Verständnis der Daten benötigten Informationen zusammengestellt werden. Hierzu zählen in der quantitativen empirischen Sozialforschung vor allem Informationen zur Entstehung der Daten insgesamt (Studiendesign, Methoden, Messinstrument) und zur Bedeutung der einzelnen Variablen des Datensatzes. Zentral für den Erhalt der dauerhaften Interpretierbarkeit sind damit insbesondere eher klassische Verfahren zur Erschließung und Dokumentation. Technisch kommen hingegen Strategien wie etwa Emulation (Nachbildung der technischen Umgebung der Objekte) oder Migration (Dateiformate, Speichermedien, technische Umgebungen) zum Einsatz.

Mit Blick auf die Nutzbarkeit von Forschungsdaten – als eine besondere Form digitaler Objekte – kommt der *Kuratierung* und Pflege dieser Daten eine besondere Rolle zu. Digitales Kuratieren (*digital curation*) ist ein Konzept, das sich in weiten Teilen mit Konzepten und Tätigkeiten im Bereich der digitalen Langzeitarchivierung überschneidet, wie die Definitionen in Schaukasten 7.1 verdeutlichen.

Schaukasten 7.1: Zwei Beispielformulierungen von *digital curation*

- „Digital curation is all about maintaining and adding value to a trusted body of digital information for future and current use; specifically, the active management and appraisal of data over the entire life cycle. Digital curation builds upon the underlying concepts of digital preservation whilst emphasising opportunities for added value and knowledge through annotation and continuing resource management.“ (Pennock 2006: 1)
- „Digital curation involves maintaining, preserving and adding value to digital research data throughout its lifecycle.“ (Digital Curation Centre o.J.)

Aufgrund dieser großen inhaltlichen Überschneidungen werden die Begriffe *Kuratierung* und *Langzeitarchivierung* häufig nicht trennscharf verwendet. *Kuratierung* soll im Sinne der hier zitierten Definitionen als Oberbegriff für alle Aktivitäten im Lebenszyklus von digitalen Objekten verstanden werden, die nicht nur darauf abzielen, diese zu erhalten und nutzbar zu machen, sondern insbesondere auch, die digitalen Objekte mit Mehrwert (*added value*) zu versehen.

Dieser Mehrwert kann z.B. darin bestehen, dass die Daten mit zusätzlichen Informationen, etwa über Eigenschaften der jeweiligen politischen Systeme in international-vergleichenden Studien, versehen werden. Ein Mehrwert wird aber auch dann geschaffen, wenn die Daten selbst ergänzt werden, z.B. durch die Imputation fehlender Werte. Die Begriffe *Aufbereitung* und *Dokumentation* fallen unter diesen Oberbegriff und können als einzelne Arbeitsschritte im Rahmen der Kuratierung angesehen werden. Während die Aufbereitung sich insbesondere auf die Arbeit an den Daten selbst bezieht, werden im Rahmen der standardisierten Datendokumentation die Daten zum Zwecke der inhaltlichen Einordnung und Nachvollziehbarkeit beschrieben (vgl. Kapitel 6). Dass in diesem Kontext unterschiedliche Arten von Metadaten bei der Dokumentation, Erschließung und langfristigen Bereitstellung von Studien, Daten und Kontextinformationen eine besondere Rolle spielen, wird in Kapitel 9 thematisiert. Damit werden unter den Begriff der Kuratierung auch Tätigkeiten gefasst, die nicht erst bei der Aufnahme der Daten in Datenzentren (vgl. Abschnitt 7.4), sondern die – zumindest teilweise – bereits von den datenproduzierenden Wissenschaftlerinnen und Wissenschaftlern durchgeführt werden.

### 7.1.3 Daten teilen (*Data Sharing*)

Die Diskussion um die eingangs skizzierten Entwicklungen und Forderungen wird in der Regel unter Verwendung der Schlagworte *Data Sharing*, *Open Science* und *Open Data* geführt. Unter *Data Sharing* versteht man die Praxis, in einem Projekt erhobene und/oder analysierte Forschungsdaten Dritten mit allen zum Verständnis und zur Nutzung notwendigen Materialien unter klar definierten Bedingungen zugänglich zu machen. Dieses Teilen von Daten dient entweder dem Zweck der Replikation oder dazu, neue Forschungsfragen zu beantworten. Zudem kann die Nutzung von veröffentlichten Daten im Rahmen der Lehre die Ausbildung von Nachwuchswissenschaftlerinnen und -wissenschaftlern unterstützen. Gründe für und Vorteile von *Data Sharing* sind ausführlich diskutiert (vgl. z.B. Herb 2015; Piwowar/Vision 2013).

Abbildung 7.1: Offenheitsgrade von Daten

|             | <b>Geschützt</b>   | <b>Zugänglich</b>                          |                                    | <b>Offen</b>   |                                  |
|-------------|--|--|------------------------------------|--|----------------------------------|
| Zugang      | Unter Verschluss   | Intern                                     | Auf Anfrage für bestimmte Zwecke   | Öffentlich, für bestimmte Zwecke                     | Zugang für jeden für alle Zwecke |
| Bedingungen | Individuelle Freigabe für bestimmte, identifizierte Personen | Arbeitsvertrag und Betriebsvereinbarungen  | Nutzungsvertrag                    | Lizenz mit Einschränkung der Nutzung (z.B. CC-BY-ND) | Offene Lizenz (z.B. CC0, CC-BY)  |
| Beispiel    | Daten mit Relevanz für die nationale Sicherheit              | Nicht-anonymisierte Daten zum Krankenstand | Schwach anonymisierte Umfragedaten | Twitter Feed   | Fahrplan                         |

Quelle: Eigene Darstellung nach Open Data Institute (o.J.)

Zu beachten ist, dass *Data Sharing* – anders als die Begriffe *Open Data* und *Open Science* implizieren – nicht notwendig bedeutet, alle Forschungsergebnisse, und insbesondere die gewonnenen Forschungsdaten, unmittelbar und für jeden frei im Internet zugänglich zu machen (Herb 2015: 26). So lässt sich ein offener Zugang, verstanden als die Ermöglichung einer freien Nutzung zu allen Zwecken und für alle interessierten Personengruppen, nicht immer realisieren. Dies betrifft u.a. personenbezogene Daten, aber auch Daten, die im Zusammenhang mit gewerblichen Schutzrechten (z.B. Patenten) stehen. Weiterhin sind Daten zu berücksichtigen, deren Veröffentlichung aus Sicherheitsgründen (z.B. nationale Sicherheit) nicht möglich ist oder deren Bekanntwerden den Untersuchungsgegenstand gefährden würde. Dies kann z.B. in der Biodiversitätsforschung der Fall sein, wenn die Daten Hinweise enthalten, die zur unerwünschten Lokalisierung von seltenen Tierpopulationen oder schützenswerter Natur durch Dritte führen könnten. Dies bedeutet aber nicht, dass solche Daten anderen überhaupt nicht zugänglich gemacht werden können. Vielmehr ist es hilfreich, Zugänglichkeit nicht so sehr als binären Gegensatz von *open* vs. *closed* zu denken, sondern, wie in Abbildung 7.1 dargestellt, als Spektrum von Offenheitsgraden zu verstehen. Innerhalb dieses Spektrums existieren abgestufte Möglichkeiten, Forschungsdaten je nach Sensitivität oder unter Berücksichtigung anderer rechtlicher und organisatorischer Rahmenbedingungen unter geeigneten Zugangs- und Nutzungsbedingungen mit einer definierten Zielgruppe zu teilen.

## 7.2 Zugangswege und Lizenzen

Forschende können im Rahmen des *Data Sharing* durch die Wahl geeigneter Zugangswege zu den Daten und die Vergabe von Lizenzen steuern, wie und unter welchen Bedingungen auf die Daten zugegriffen werden darf.

Zugangswege, die auch für sozialwissenschaftliche Daten gängig sind, umfassen z.B.:

- *Freier Zugang über das Internet*: Daten und Dokumentation sind online verfügbar und können über das Internet ohne weitere Hürden zur Nutzung heruntergeladen werden;
- *Zugang nach Registrierung*: Nutzende müssen einen Account beim Anbieter der Daten erstellen und sich anmelden, um Zugang zu Daten zu erhalten;
- *Geschützter Remote-Zugang*: Besonders sensitive Daten werden – in der Regel nach Antragstellung und nach Abschluss eines Vertrags – in einer besonders geschützten Remote-Umgebung zur Analyse bereitgestellt. Dies bedeutet, dass Forschende zwar über das Internet auf die Daten zugreifen können,

dieser Zugriff aber durch technische Schutzmechanismen beschränkt wird (vgl. Schiller et al. 2017). So ist ein Herunterladen der Daten in der Regel nicht möglich und erstellte Analyseergebnisse werden ggf. von Mitarbeitenden des Datenanbieters auf Einhaltung des Datenschutzes kontrolliert, bevor sie an die Forschenden weitergegeben werden;

- *Vor-Ort-Zugang*: Forschungsdaten sind nach Antragstellung und Vertragsschluss in einer geschützten Umgebung lokal beim Datenanbieter zugänglich. Auch hier werden Analyseergebnisse in der Regel kontrolliert, bevor sie ausgehändigt werden.

Zu beachten sind bei der Wahl eines Zugangswegs für Forschungsdaten – sofern relevant – die Regelungen des Datenschutzes sowie Zusicherungen und Angaben, die Studienteilnehmenden im Zusammenhang mit der informierten Einwilligung gegeben wurden. Zudem ist zu beachten, dass (die Entscheidung über) eine Veröffentlichung und Modalitäten der Nutzung der Daten nur von Inhaberinnen und Inhabern der Verwertungsrechte erfolgen kann. Dies kann insbesondere im Fall von Sekundäranalysen zu Problemen führen, wenn in der eigenen Forschung bereits existierende Datenbestände Dritter genutzt werden. Hier ist es in der Regel notwendig, Rechteinhaber/-innen zu kontaktieren und eine explizite Genehmigung für eine Veröffentlichung einzuholen.

Innerhalb des formalen Rahmens, der durch die oben genannten Zugangswege geschaffen wird, können zusätzliche Nutzungsbedingungen definiert werden. Diese legen fest, durch wen und für welche Nutzungszwecke Forschungsdaten zugänglich sein sollen. Solche Nutzungsbedingungen werden auch als Lizenzen bezeichnet. So schreibt Kreutzer (2016: 19):

Eine Lizenz ist eine Nutzungserlaubnis für Handlungen, die ohne Zustimmung nicht erlaubt wären. Ob eine Lizenz ein Vertrag oder ein einfaches, einseitiges Versprechen ist, ist von Land zu Land verschieden. Die Auswirkungen sind hingegen dieselben: Die Lizenz ist eine rechtlich gültige Vereinbarung, die die Verwendung eines bestimmten Werkes regelt. Verwendungen, die nicht von der Lizenz abgedeckt sind oder die gegen die Lizenzpflichten verstoßen, sind widerrechtliche Handlungen, die rechtliche Folgen nach sich ziehen können.

Lizenzen sind ein hilfreiches Instrument, um die Nutzung von Daten durch Dritte zu regeln. Sie erlauben es einerseits – im Sinne eines möglichst offenen Zugangs zu den Ergebnissen öffentlich geförderter Forschung –, Regeln aufzustellen, die im Vergleich zum deutschen Urheberrecht die Nutzung der Daten erleichtern und Hürden abbauen. Andererseits schaffen Lizenzen Transparenz über erlaubte Nutzungsformen, da sie für Nutzende oft leichter verständlich sind als das Urheberrecht. Dies gilt umso mehr, wenn Lizenztexte nicht frei formuliert werden, sondern standardisierte und möglichst weit verbreitete Lizenzen verwendet werden, wie beispielsweise die in Schaukasten 7.2 dargestellten Creative Commons-Lizenzen.

Bei der Nutzung von Creative Commons (CC) und anderen Standardlizenzen für Forschungsdaten ist zu beachten, dass diese Lizenzen erforderliche oder erwünschte Zugangs- und Nutzungsszenarien nicht immer abbilden können. Zum Beispiel können Daten, die unter besonderem Schutz stehen und deshalb nur in einer geschützten technischen Umgebung verarbeitet werden dürfen, nicht unter eine CC-Lizenz gestellt werden, da diese immer auch die Weitergabe der betreffenden Daten an Dritte erlaubt. Ebenso ist die Beschränkung auf eine rein wissenschaftliche Nutzung der Daten mit Hilfe von CC-Lizenzen nicht darstellbar.

Die Klärung rechtlicher Aspekte der Nutzung von Forschungsdaten durch Dritte ist ein wichtiger Schritt in der Entscheidung, ob und wie diese Daten anderen zur Nutzung bereitgestellt werden können und sollen. Ist dies erfolgt, können weitere Schritte zur Auswahl einer geeigneten Lösung für die Sicherung sowie ggf. Archivierung und Veröffentlichung der Daten erfolgen. Hierzu geben die folgenden Abschnitte einen Überblick darüber, welche Arten von unterschiedlichen Diensten es gibt, welche Leistungen sie erbringen und wie Forschende für ihre Bedürfnisse geeignete Angebote identifizieren können.

## Schaukasten 7.2: Creative-Commons-Lizenzen

- Creative-Commons-(CC)-Lizenzen beruhen auf vier definierten Nutzungsbedingungen, die sich zu insgesamt sechs Standardlizenzen kombinieren lassen:
  - Namensnennung (attribution - *BY*),
  - Weitergabe unter gleichen Bedingungen (share alike – *SA*),
  - nicht-kommerziell (non-commercial – *NC*),
  - keine Bearbeitungen (no derivatives – *ND*).

Alle Kombinationen erfordern die Namensnennung der Urheberin oder des Urhebers des Werks, z.B. CC-BY-SA oder CC-BY-NC-SA (vgl. Creative-Commons - Mehr über die Lizenzen o.J.).
- Für alle CC-Standardlizenzen gibt es eine kurze Alltagssprachliche Zusammenfassung, die es den Lizenzgeberinnen und -gebern sowie den Nutzenden ermöglicht, sich schnell über den Umfang einer Lizenz zu informieren. Verbindlich für die Nutzung sind aber die ausformulierten und juristisch geprüften Lizenztexte.
- CC-Lizenzen sind prinzipiell für alle Arten von urheberrechtlich geschützten Werken nutzbar – von Bildern über Musik bis zu wissenschaftlichen Textpublikationen und Forschungsdaten. Zu beachten ist aber, dass eine Lizenz nur vom Inhaber/der Inhaberin der sogenannten Verwertungsrechte vergeben werden kann. Dies sind in der Regel die Urheberinnen und Urheber oder ihre Arbeitgeber, sofern ein Werk als Teil eines Beschäftigungsverhältnisses geschaffen wurde.
- Wollen Forschende ihre Daten ohne jegliche Einschränkung der Nutzung anderen zur Verfügung stellen, können sie auch eine sogenannte CC0-Lizenz vergeben. Eine kurze Erläuterung der Besonderheiten dieser Lizenz geben Steinhilber und Pachali (2017). Auch wenn die Namensnennung des Urhebers oder der Urheberin hier keine Nutzungsbedingung mehr darstellt, ist das Zitieren aller verwendeten Quellen im wissenschaftlichen Kontext zur Wahrung der guten wissenschaftlichen Praxis selbstverständlich trotzdem erforderlich.

Quelle: Eigene Darstellung mit Verweis auf die genannten Quellen

## 7.3 Bereitstellung und Veröffentlichung von Daten nach Projektende

Sollen Forschungsdaten nach der unmittelbaren Projektphase anderen zur Nachnutzung bereitgestellt werden, sollte dies idealerweise von Anfang an eingeplant werden. So kann der Aufwand, der ggf. am Projektende bei der Übergabe an ein Repositorium zur Veröffentlichung der Daten entsteht, teils drastisch gesenkt werden. Zusätzlich profitieren während der Projektlaufzeit alle Beteiligten von gut gepflegten und dokumentierten Daten.

### 7.3.1 Aufbewahrung zur Sicherung guter wissenschaftlicher Praxis

Selbst wenn erzeugte Forschungsdaten nach Projektende nicht veröffentlicht werden sollen oder dürfen, gilt es, die Minimalanforderungen für die Sicherung und Aufbewahrung von Forschungsdaten, wie sie sich regelmäßig in den Regeln guter wissenschaftlicher Praxis finden, im Blick zu halten. So empfiehlt beispielsweise die Deutsche Forschungsgemeinschaft (DFG 2013), an der sich viele institutionelle Regeln orientieren, Daten mindestens zehn Jahre aufzubewahren. Dies soll sicherstellen, dass Daten auf Nachfrage zu Replikationszwecken herausgegeben werden können. Auch wenn diese und ähnliche Empfehlungen (vgl. z.B. Leibniz-Gemeinschaft 2015; ALLEA 2017) das Thema Dokumentation nur rudimentär aufgreifen, sollten Forschende Vorkehrungen treffen, damit im Fall der Fälle auf die richtigen Versionen zugegriffen werden kann und auch noch nachvollziehbar ist, wie diese entstanden und zu interpretieren sind.

Die wahrscheinlich nach wie vor häufigste Form des Umgangs mit Forschungsdaten nach der Beendigung eines Projektes ist die Speicherung auf eigenen Computern, USB-Sticks oder

DVDs (vgl. z.B. Kindling/Schirmbacher/Simukovic 2013: 54; Kuipers/van der Hoeven 2009: 32). Sie ist gleichzeitig die am wenigsten geeignete, um die vorgenannten Anforderungen zu erfüllen. Solche Lösungen sind meist der Tatsache geschuldet, dass die oben angesprochenen Fragen zum Nachnutzungsszenario nicht ausreichend bzw. frühzeitig adressiert wurden. Typischerweise steht am Projektende keine Zeit zur Verfügung, sich mit Organisation, Aufbereitung, Dokumentation und Veröffentlichung der Daten zu befassen. Vielmehr müssen Abschlussberichte geschrieben und neue Projekte beantragt werden. Nach Projektende lösen sich Projektteams auf, Mitarbeitende übernehmen neue Aufgaben oder wechseln zu einer anderen Institution. All dies führt nicht selten dazu, dass Forschungsdaten selbst den unmittelbar an der Erhebung beteiligten Wissenschaftlerinnen und Wissenschaftlern nicht mehr in geeigneter Form zur Verfügung stehen, um zu einem späteren Zeitpunkt Analyseergebnisse zu überprüfen oder weiterführende Fragestellungen zu bearbeiten. Aber nicht nur die Speicherung auf dafür nicht geeigneten Datenträgern bzw. in unsicheren Umgebungen birgt Risiken.

Selbst eine planvollere Herangehensweise, die eine professionelle und redundante Datenhaltung beinhaltet, ist nicht unbedingt eine Garantie dafür, dass Daten über einen längeren Zeitraum nutz- und interpretierbar bleiben. Unsicherheit bzgl. vorhandener Versionen, undokumentierte Bereinigungen, Aufbereitungen oder sonstige Veränderungen an den Daten, unzureichende Dokumentation der Datenerhebung oder der Stichprobenziehung kompromittieren den Wert von Daten für eine spätere Verwendung. Einen oft übersehenen Risikofaktor stellen die gewählten Dateiformate dar. So können Dateien komplett unlesbar werden oder sich Einschränkungen in der Nutzbarkeit der Daten ergeben, wenn die zur Erstellung ursprünglich genutzte Software nicht mehr oder nur in neueren Versionen zur Verfügung steht. Insbesondere wenn längere Zeiträume der Nachnutzung angestrebt werden, bedarf es daher eines systematischen Ansatzes zur digitalen Langzeitarchivierung, wie oben bereits erwähnt.

### 7.3.2 Optionen zur Veröffentlichung und Archivierung von Forschungsdaten

Wenn Forschungsdaten veröffentlicht werden sollen, bieten sich hierfür verschiedene Wege oder Formen an. Mittlerweile existiert eine Vielzahl von Diensten, die Forschende bei der Sicherung, Aufwertung und Veröffentlichung von Forschungsdaten unterstützen. Diese unterscheiden sich jedoch z.T. erheblich, z.B. hinsichtlich Art und Umfang der angebotenen Dienstleistungen, der fachlichen Abdeckung oder auch der Art und Weise, wie Daten Dritten zur Nachnutzung zur Verfügung gestellt werden. Neben generischen Angeboten, die offen für Daten aus allen Fachgebieten sind, finden sich solche, die auf bestimmte Disziplinen, Themen oder Datentypen spezialisiert sind (s. Schaukasten 7.3 für Beispiele).

Institutionelle Repositorien konzentrieren sich in der Regel auf Forschungsausgaben der eigenen Einrichtung, während disziplin- oder domänenspezifische Forschungsdateninfrastrukturen eher auf nationaler oder internationaler Ebene arbeiten, dafür aber nicht allen Communities oder Datentypen offenstehen. Neben der eigenständigen Veröffentlichung der Forschungsdaten in einem Repositorium oder Datenzentrum, besteht auch die Möglichkeit, ein *Data Paper* in einem Datenjournal zu veröffentlichen (s.u.). Diese konzentrieren sich im Gegensatz zu klassischen Fachzeitschriften auf die Publikation von Forschungsdaten bzw. auf Beschreibungen dieser Daten. Darüber hinaus werden Forschungsdaten aber auch als Ergänzung zu einer wissenschaftlichen (Text-)Publikation veröffentlicht, z.B. als *Supplementary Material*.

Welcher Weg nun für die Sicherung, Archivierung und Veröffentlichung der eigenen Forschungsdaten am geeignetsten ist, kann pauschal nicht beantwortet werden, sondern hängt

von einer Vielzahl von Faktoren ab. Einige Hinweise werden im folgenden Abschnitt gegeben.

Unter *Repositorien* werden in diesem Kontext über das Internet zugängliche Plattformen zur Sicherung und Veröffentlichung von Forschungsdaten verstanden. Datenproduzentinnen und -produzenten können Forschungsdaten über ein Online-Formular in das Repositorium laden und mit mehr oder weniger reichhaltigen Metadaten beschreiben. Im Zuge der Veröffentlichung werden die Daten mit einem persistenten Identifikator versehen, sodass sie dauerhaft auffindbar, referenzierbar und zitierbar sind, wie in Kapitel 10 behandelt. Neben dem Vorteil einer systematischen Sicherung der eigenen Daten bieten Repositorien eine sehr gute Möglichkeit, Daten anderen Personen zur Verfügung zu stellen. Idealerweise können Forschende dabei selbst entscheiden, wem und unter welchen Bedingungen ihre Daten zur Verfügung gestellt werden. Dies geschieht sinnvollerweise durch die Verwendung standardisierter und – sofern möglich – offener Lizenzen (s. Abschnitt 7.2).

| Schaukasten 7.3: Angebote zur Sicherung und Veröffentlichung von Forschungsdaten (Beispiele) |  |   |
|--|--|---|
| <i>Art des Dienstes</i>  | <i>Hinweise</i>  | <i>Beispiele</i>  |
| Fachübergreifend, nicht institutionell gebunden  | Geschäftsbedingungen prüfen – z.B. auf (unerwünschte) Gewährung von Rechten zur Datennutzung, Aspekte der Datensicherheit und des Datenschutzes, garantierte Haltefristen.<br>Für die Inanspruchnahme der Dienste zur Veröffentlichung können (einmalige oder wiederkehrende) Kosten anfallen. | Kommerzieller Träger <ul style="list-style-type: none"> <li>• Figshare</li> <li>• Mendelej Data</li> </ul> Öffentlich geförderter Träger <ul style="list-style-type: none"> <li>• Zenodo</li> <li>• RADAR (kostenpflichtig)</li> </ul>  |
| Institutionelles Repositorium  | Einreichung von Daten in der Regel nur für Institutionsangehörige  | <ul style="list-style-type: none"> <li>• heiDATA, Universität Heidelberg</li> <li>• Forschungsdatenzentrum des Robert Koch-Instituts</li> </ul>   |
| Disziplinspezifische Angebote  | Frühzeitig Anforderungen für eine Einreichung zur Archivierung und Veröffentlichung in Erfahrung bringen (z.B. in Bezug auf Datenformate, Dokumentation).<br>Für die Inanspruchnahme der Dienste zur Veröffentlichung können (einmalige oder wiederkehrende) Kosten anfallen.                  | <ul style="list-style-type: none"> <li>• PsychData, Leibniz-Zentrum für Psychologische Information und Dokumentation, für Daten der Psychologie</li> <li>• Qualiservice Bremen für qualitative sozialwissenschaftliche Daten</li> <li>• GESIS Datenarchiv für Sozialwissenschaften für quantitative sozialwissenschaftliche Daten</li> <li>• Datenservicezentrum Betriebs- und Organisationsdaten, Universität Bielefeld</li> </ul> |

Quelle: Eigene Darstellung

In den letzten Jahren wurden insbesondere an Hochschulen Repositorien für Forschungsdaten aufgebaut (teilweise auf Grundlage der bereits existierenden Dienste für Open-Access-Publikationen), um Forschende an den jeweiligen Einrichtungen beim Umgang mit den veränderten Anforderungen in Bezug auf Forschungsdaten zu unterstützen. Meist werden zusätzliche Informations- und Unterstützungsangebote rund um das Thema Forschungsdatenmanagement und die Veröffentlichung der Daten im hauseigenen Dienst angeboten.<sup>1</sup>

Zu den Vorteilen solch institutioneller Dienste zählen ihre räumliche und organisatorische Nähe und die auf die jeweiligen institutionsspezifischen Bedürfnisse, Regeln und

1 Häufig werden solche Beratungsangebote an Hochschulen und außeruniversitären Forschungseinrichtungen gemeinsam mit einer Forschungsdatenleitlinie etabliert. Daher stellt die regelmäßig aktualisierte Übersicht über institutionelle Policies im Wiki von [www.forschungsdaten.org](http://www.forschungsdaten.org) einen guten Ausgangspunkt dafür dar, sich einen Überblick über existierende Beratungsangebote zu verschaffen.

Policies angepassten Angebote. So stellen institutionelle Repositorien einerseits eine sehr gute Anlaufstelle für Forschende dar. Andererseits ermöglichen sie es der betreibenden Einrichtung, den eigenen Forschungsoutput gut organisiert und sichtbar darzustellen, insbesondere wenn sie über geeignete Schnittstellen an ein Forschungsinformationssystem zur Erfassung solcher Forschungsoutputs angebunden sind.

Ungeachtet der Unterschiede in der technischen oder organisatorischen Ausgestaltung der jeweiligen institutionellen Repositorien ist diesen zumindest im hochschulischen Kontext typischerweise gemein, dass sie Daten aus allen oder zumindest mehreren Disziplinen offenstehen. Dies ist einerseits ein Vorteil, da derartige niedrigschwellige Angebote die Möglichkeit bieten, vielfältige Formen von Forschungsdaten zu archivieren und zu veröffentlichen, um so beispielsweise entsprechenden Anforderungen von Förderern zu begegnen. Andererseits folgt aber daraus auch, dass sie nicht auf Besonderheiten der jeweiligen Disziplinen respektive Daten ausgerichtet werden können. So müssen beispielsweise eher generische Metadatenschemata zur Beschreibung der Forschungsdaten eingesetzt werden, die sich häufig auf bibliographische Angaben beschränken und daher nur eingeschränkt geeignet sind, Spezifika der jeweiligen Daten und Disziplinen abzubilden. Auch haben die meist sehr geringen Anforderungen hinsichtlich Dokumentation, Aufbereitung und technischer Aspekte (z.B. Formate) Auswirkungen auf die Nutzbarkeit der Daten. So kann nicht wirklich sichergestellt werden, dass die veröffentlichten Datensätze langfristig nutzbar sein werden bzw. dass ihre Analysepotentiale vollständig ausgenutzt werden können. Häufig sind zudem die Möglichkeiten einer fachspezifischen Beratung der Forschenden aufgrund knapper Ressourcen eingeschränkt. Gerade in dem für die Sozialwissenschaften wichtigen Themenkomplex Datenschutz, Anonymisierung und informierte Einwilligung fehlen oft entsprechende Beratungsangebote. Auch haben institutionelle, disziplinübergreifende Repositorien in der Regel eine geringere Sichtbarkeit innerhalb der Fachcommunity, sodass die Daten von der Zielgruppe kaum gefunden werden können.

Insbesondere in den letztgenannten Bereichen liegen die Stärken von fachspezifischen Angeboten (s. Schaukasten 7.3 für Beispiele). Auch in diesem Segment gibt es allerdings eine große Heterogenität in Bezug auf Art, Umfang und Struktur der jeweiligen Angebote. Die Bandbreite reicht von einfachen Angeboten zur Sicherung und Veröffentlichung von Forschungsdaten mittels Repositorien bis hin zu komplexen Forschungsdateninfrastrukturen, die, teils auf internationaler Ebene, vielfältige Dienste rund um die Aufbereitung, Dokumentation, Registrierung und Langzeitarchivierung anbieten sowie umfangreiche Informations-, Beratungs- und Schulungsangebote bereithalten.

In den Sozialwissenschaften existieren solche Dateninfrastrukturen schon recht lange (Mochmann 2008). Der Aufbau begann bereits kurz nach dem Zweiten Weltkrieg und entwickelte in den 1960er Jahren – u.a. mit der Etablierung des Zentralarchivs für Empirische Sozialforschung, dem heutigen GESIS Datenarchiv für Sozialwissenschaften – weltweit eine gewisse Dynamik. Allein in Europa sind derzeit sozialwissenschaftliche Datenserviceeinrichtungen aus 34 Ländern im Consortium of European Social Science Data Archives (CESSDA) organisiert, wovon gegenwärtig 17 dieser Einrichtungen im Rahmen eines sogenannten European Research Infrastructure Consortium (ERIC) eine gemeinsame europäische Dateninfrastruktur aufbauen. In den deutschen Sozialwissenschaften hat sich darüber hinaus eine Reihe spezialisierter Forschungsdatenzentren<sup>2</sup> herausgebildet, die Datenservices ausschließlich für bestimmte Forschungsdaten anbieten – meist für Daten, die an der jeweiligen Institution selbst erhoben werden – und in der Regel nicht für die Aufnahme von externen Daten zur Verfügung stehen. Bei diesen Zentren handelt es sich meist um kleinere Einheiten, die bei der datenerzeugenden oder -haltenden Organisation angesiedelt sind und einen

---

2 Diese sind in die Forschungsdateninfrastruktur des Rats für Sozial- und Wirtschaftsdaten eingebunden.

thematisch engen Fokus haben, um eine hohe fachwissenschaftliche Qualität der Datenhandhabung zu gewährleisten. Einen Überblick über die Struktur und Entwicklung sozialwissenschaftlicher Infrastrukturen bieten Quandt und Mauer (2012), Mauer (2012) sowie Scheuch (2003).

Darüber hinaus haben sich in jüngster Zeit vermehrt auch sogenannte Datenjournale (*Data Journals*) etabliert, die sich im Gegensatz zu klassischen Fachzeitschriften auf die Publikation von Forschungsdaten bzw. Beschreibungen von Forschungsdaten konzentrieren.<sup>3</sup> Die Daten selbst werden meist in kooperierenden Repositorien veröffentlicht. In der Regel ist eine solche ‚Datenpublikation‘ an ein Begutachtungsverfahren gekoppelt. Neben klassischen Peer-Reviews, die beispielsweise durch eine Begutachtung von Datenspezialistinnen und -spezialisten ergänzt werden (wie beispielsweise beim *Research Data Journal for the Humanities and Social Sciences*), werden auch innovative Verfahren zur Begutachtung eingesetzt. So kommt im Fall des *Earth System Science Data* (ESSD) ein mehrstufiges *Interactive-Public-Peer-Review*-Verfahren zum Einsatz. Datenjournale tragen in der nach wie vor vorherrschenden Kultur von *publish or perish* dem Umstand Rechnung, dass in der Vergangenheit die erhebliche Arbeit, die in die Aufbereitung von Daten für die Forschung und insbesondere für eine Nachnutzung durch Dritte fließt, nicht ausreichend gewürdigt oder gar als Forschungsleistung anerkannt wurde. Indem *Data Papers* als referierte Publikation anerkannt werden können, schaffen sie einen zusätzlichen Anreiz zum Teilen und Veröffentlichenden von Forschungsdaten.

Ebenso ist es gängig, die einem wissenschaftlichen Artikel zugrunde liegenden Forschungsdaten als *Supplementary Material* der Textpublikation unmittelbar beizufügen, d.h. sie auf den Webseiten der Zeitschrift bzw. des Verlags zu veröffentlichen. Diese Supplemente erhalten häufig keinen eigenen persistenten Identifikator und fallen unter die gleichen Zugriffsschranken wie ggf. der Artikel selbst (vgl. National Information Standards Organization 2013). Zudem ist in diesem Fall zu beachten, dass Autorinnen und Autoren möglicherweise vertragliche Vereinbarungen mit dem Verlag schließen (*Autorenvertrag*), die diesem ggf. umfassende Verwertungsrechte an den Daten einräumen. Dies kann Implikationen für die Verwendung der Daten in der eigenen Forschung haben. Auch müssen Forschende sich bei der Frage der langfristigen Sicherung und Bereitstellung der Daten auf den Verlag verlassen (vgl. Reilly et al. 2011: 47). Eine Veröffentlichung in einem Repository und Verknüpfung mit dem Artikel über einen persistenten Identifikator ist daher vorzuziehen. Wenn dies nicht möglich oder erwünscht ist, sollten Forschende darauf achten, dass sie dem Verlag nur ein einfaches Nutzungsrecht an den Daten übertragen und stattdessen ihre Daten soweit als möglich unter einer offenen Lizenz veröffentlichen.

Im folgenden Abschnitt sollen die Abläufe einer Veröffentlichung von Forschungsdaten in einem disziplinspezifischen oder institutionellen Datenzentrum skizziert werden. Dabei wird vor allem auf für sozialwissenschaftliche Forschungsdaten relevante Aspekte eingegangen.

## 7.4 Überblick: Was machen Datenzentren mit meinen Forschungsdaten?

Disziplinspezifische Datenzentren bieten forschungsdatenbezogene Services an, die auf die besonderen Bedarfe einer Disziplin oder Community zugeschnitten sind. Sie verfügen in der

---

3 Eine Übersicht über Datenjournale findet sich bei [www.forschungsdaten.org](http://www.forschungsdaten.org).

Regel über die dafür notwendige fachwissenschaftliche Expertise, denn nur so können die Spezifika der jeweiligen Forschungsprozesse und der dort entstehenden bzw. verwendeten Forschungsdaten in entsprechenden Services abgebildet werden.

Im GESIS Datenarchiv durchlaufen eingehende Forschungsdaten und Dokumentationen einen mehrstufigen Prozess, der letztlich darauf ausgerichtet ist, die Nachvollziehbarkeit und das Analysepotential der Daten zu erhöhen sowie ihre Nachnutzbarkeit langfristig zu sichern. Die einzelnen Schritte dieses Prozesses können entsprechend den Bedarfen der Forschenden bzw. ihrer Daten in Umfang und Tiefe unterschiedlich ausgestaltet werden.

#### 7.4.1 Vorbereitung der Datenübergabe

Idealerweise klären Forschende und Dateninfrastruktur vor der eigentlichen Übergabe von Forschungsdaten zunächst Details bzgl. Art, Umfang und Zustand der jeweiligen Daten und besprechen, welche Aufbereitungs- und Dokumentationsziele angestrebt werden. Auch die Frage, auf welchem Weg und unter welchen Bedingungen die Daten zur Nachnutzung bereitgestellt werden sollen bzw. können, sollte vor der Übergabe geklärt werden. Dies kann im Gespräch mit Archivmitarbeitenden erfolgen oder auf Grundlage der auf der Webseite bereitgestellten Informationen entschieden werden. Diese Absprachen werden in der Regel in einer sogenannten Archivierungsvereinbarung festgehalten (vgl. Schaukasten 7.4). In der Vereinbarung wird auch die für eine Archivierung und Weitergabe notwendige Übertragung von Nutzungsrechten an das Archiv definiert. Diese Vorbereitung sollte möglichst früh im Lebenszyklus einer Studie ansetzen, sodass die Archivierung und damit insbesondere eine Nachnutzung durch Dritte möglichst reibungslos ablaufen können (s. Kapitel 3).

##### Schaukasten 7.4: Typische Inhalte einer Archivierungsvereinbarung

- Übertragung von Rechten zur Speicherung, Vervielfältigung und Verbreitung der Forschungsdaten unter den vereinbarten Bedingungen
- Festlegung der Zugangsbedingungen zu den Daten
- Bestätigung des Datengebers/der Datengeberin, dass die Rechte zur Veröffentlichung der Daten bei ihm/ihr liegen
- Rechtsnachfolge: Was geschieht beim Ableben des Datengebers/der Datengeberin?
- Verpflichtung zur Einhaltung der Regelungen des Datenschutzes durch Datengeber/-innen und das Archiv

Quelle: Eigene Darstellung

#### 7.4.2 Datenübergabe und Übernahme in das Archiv

In der Regel erfolgt die Datenübergabe an das GESIS Datenarchiv entweder in Absprache mit den verantwortlichen Mitarbeitenden über ein Tool zur sicheren Datenübertragung oder über ein Formular. In diesem ergänzen Forschende beschreibende Metadaten zu den eingereichten Forschungsdaten und laden die Daten dann zusammen mit allen erforderlichen Dokumenten – üblicherweise Codebuch, Erhebungsinstrument und ggf. Syntaxen – hoch.

Bei der sich anschließenden Aufnahme der Daten in das Archiv werden diese durch erfahrene Kuratorinnen und Kuratoren geprüft. Eine solche standardisierte Eingangskontrolle umfasst beispielsweise Prüfungen der folgenden Aspekte:

- Technische Kontrollen der Dateien: Liegen die Daten in einem akzeptierten Format vor? Sind die Dateien lesbar und virenfrei?
- Vollständigkeit und Nutzbarkeit: Sind Daten, Messinstrument, Methodenbericht, Datendokumentation und weitere Dokumente wie vereinbart und vollständig übergeben worden? Beziehen sich alle

übergebenen Materialien logisch aufeinander? Insbesondere: Passen Erhebungsinstrument, Datensatz und Datendokumentation zusammen?

- Konsistenz: Gibt es Werte außerhalb des zulässigen Bereichs (*wild codes*)? Sind fehlende Werte definiert und wenn ja, wie? Stimmen die Daten mit der im Fragenbogen vorgegebenen Filterführung überein (*question routing*)? Gibt es widersprüchliche Merkmalskombinationen?
- Datenschutz: Da es sich bei den durch GESIS archivierten Studien vorrangig um Mikrodaten handelt, werden sie darüber hinaus auch auf datenschutzrechtlich relevante Aspekte untersucht.

Bei der Eingangskontrolle festgestellte Inkonsistenzen oder sonstige Fehler werden dokumentiert und zurückgemeldet. In Abhängigkeit vom Ergebnis der Eingangskontrolle und einem zuvor festgelegten Aufbereitungs- und Dokumentationsziel folgen der Eingangskontrolle weitere Arbeitsschritte, wie beispielsweise die Konvertierung der übergebenen Dateien in definierte Langfristsicherungsformate und deren Überführung in den zentralen Archivspeicher.

#### 7.4.3 *Aufbereitung und Dokumentation der Daten*

Ein wichtiger Bestandteil der Kuratierung von Forschungsdaten besteht darin, diese so aufzubereiten und zu dokumentieren, dass sie durch Dritte gefunden und genutzt werden können. Der dabei vom Datenarchiv übernommene Umfang dieser Arbeiten ist einerseits abhängig vom jeweiligen Ausgangszustand der Daten und andererseits vom konkreten Aufbereitungs- und Publikationsziel der jeweiligen Studie. Das Spektrum reicht von reinen Empfehlungen, wie die betreffenden Daten verbessert oder aufgewertet werden können, über die Erstellung von Studienbeschreibungen, die in standardisierter Form (kompatibel zur Metadatenpezifikation der internationalen Data Documentation Initiative, DDI) inhaltliche, methodische und technische Charakteristika einer Studie spezifizieren, bis hin zur Übernahme umfangreicher und komplexer Aufbereitungs- und Dokumentationsarbeiten durch das Archiv. Diese umfassen beispielsweise:

- weitergehende Standardisierung und Aufbereitung der Daten,
- Harmonisierung und Integration von Daten zu komplexen zeit- und/oder ländervergleichenden Datensätzen,
- standardisierte Dokumentation des Erhebungsinstruments, inkl. Filter- und Intervieweranweisungen,
- Erweiterung der Dokumentation um Metadaten auf Studienebene.

Dieser umfangreiche Bestand an strukturierten Metadaten kann zum einen über verschiedene Datenportale und Recherchesysteme zur Verfügung gestellt werden. Zum anderen dient er auch dazu, begleitende Dokumentationen etwa in der Form von Codebüchern oder Methodenberichten zu erstellen.

#### 7.4.4 *Bereitstellung von Daten und Dokumentationen*

Zentrales Ziel der vorausgegangenen Tätigkeiten ist die Bereitstellung der Forschungsdaten für die Nachnutzung; sei es, um publizierte Forschungsergebnisse nachzuvollziehen oder aber – und das gilt für die überwältigende Anzahl der Fälle – um neue Forschungsfragen mit Hilfe dieser Daten zu beantworten (vgl. Kapitel 8). In der Regel werden Daten und weitere Materialien, wie etwa Fragebögen oder Codebücher, in Portalen zum direkten Download angeboten. Das GESIS Datenarchiv unterhält wie alle Archive der CESSDA-Infrastruktur einen Online-Katalog (Datenbestandskatalog, DBK), der als zentraler Zugangspunkt Informationen zu allen archivierten Studien bereitstellt. Neben bibliographischen Angaben werden dort insbesondere Inhalte und methodische Aspekte beschrieben, wie etwa die

Stichprobenziehung oder die in der Studie eingesetzten Erhebungsverfahren. Darüber hinaus finden sich auch Hinweise auf Veröffentlichungen und die Versionsgeschichte sowie aktuell bekannte Fehler in den Daten. Studien, die beispielsweise aus datenschutzrechtlichen Gründen nur nach Abschluss eines gesonderten Nutzungsvertrags zugänglich gemacht werden, können über ein Warenkorbsystem bestellt werden. Besonders schützenswerte Daten werden über das Secure Data Center (SDC) in eigens dafür eingerichteten Räumen – sogenannten *Safe Rooms* – zur Vor-Ort-Nutzung bereitgestellt. Eine ausführliche Darstellung von Services des GESIS Datenarchivs im Bereich der Aufbereitung, Dokumentation und Bereitstellung von Daten findet sich in Recker, Zenk-Möltgen und Mauer (2017).

Der Umfang und die konkrete Ausgestaltung der hier am Fallbeispiel GESIS geschilderten Schritte können sich von Infrastruktur zu Infrastruktur unterscheiden und somit als ein Kriterium im Prozess der Auswahl eines geeigneten Repositoriums dienen. Im folgenden Abschnitt sollen nun abschließend noch weitere Hinweise zur kriteriengeleiteten Auswahl eines geeigneten Dienstes gegeben werden.

## 7.5 Auswahl eines geeigneten Repositoriums

Das Angebot an Diensten zur Sicherung, Archivierung und Veröffentlichung von Forschungsdaten ist mittlerweile groß und entsprechend unübersichtlich. Forschende stehen damit vor der Schwierigkeit, passende Angebote für ihre spezifischen Anforderungen zu identifizieren, zu bewerten und eine – idealerweise kriteriengestützte – Auswahlentscheidung zu treffen. Bei der Identifikation geeigneter Angebote können Forschende folgendermaßen vorgehen:

Viele Hochschulen und Forschungseinrichtungen unterhalten inzwischen institutionelle Angebote für die Sicherung und Veröffentlichung der an der eigenen Einrichtung generierten Daten. Alternativ, insbesondere wenn das institutseigene Repositorium eine disziplinübergreifende, generische Ausrichtung hat, werden oftmals Empfehlungen für fachspezifische Angebote ausgesprochen. Hier können beispielsweise Bibliotheken und/oder besonders Servicestellen für Forschungsdatenmanagement weiterhelfen, wie sie bereits an einer Reihe deutscher Hochschulen bestehen.

Wenn die eigene Institution keine Orientierung oder Hilfestellung bietet, können auch wissenschaftliche Verlage als Anlaufstellen dienen. So empfehlen PLOS One oder Nature geeignete Repositorien für die Publikation von Daten zu bei ihnen veröffentlichten Artikeln.

Eine weitere Möglichkeit, einen geeigneten Dienst für die Archivierung und/oder Bereitstellung der erhobenen Daten zu identifizieren, sind Onlineverzeichnisse für Forschungsdatenrepositorien. Nationale Forschungsinfrastrukturen können im Informationsportal RISources der DFG recherchiert werden. Internationale Angebote sind im Onlineverzeichnis Re3Data verzeichnet, wo sie mit Hilfe eines eigens entwickelten Metadatenschemas (Rücknagel et al. 2015) beschrieben werden und über eine facettierte Suche auffindbar sind. Um in Re3data einen geeigneten Dienst für die Veröffentlichung von Daten zu identifizieren, sind – neben der Auswahl der Disziplin (*subject*) und/oder des Datentyps (*content type*) – vor allem die in Schaukasten 7.5 erläuterten Felder von Bedeutung.

| Schaukasten 7.5: Ausgewählte Metadatenfelder in der Re3Data-Suche                        |   |
|--|---|
| <i>Feld</i>  | <i>Erläuterung</i>  |
| Data upload (Data upload restrictions)   | Die Felder geben Auskunft darüber, wer in einem Dienst Daten veröffentlichen kann. Nur Dienste mit dem Merkmal <i>Data upload: Open</i> stehen grundsätzlich allen Forschenden für die Veröffentlichung von Daten offen. Doch auch Dienste mit dem Merkmal <i>Data upload: Restricted</i> können häufig nach Registrierung ( <i>Data upload restriction: Registration</i> ) ohne weitere Einschränkungen zur Veröffentlichung genutzt werden. |
| Data access (Data access restrictions)<br>Database access (Database access restrictions) | Die Felder enthalten Informationen darüber, unter welchen Bedingungen die Daten und das Repositorium für Nutzende zugänglich sind. Liegen Zugangsbeschränkungen (z.B. Zugang nur nach Registrierung) vor, werden sie im entsprechenden Feld näher beschrieben.  |
| Certificates   | Das Feld enthält Informationen über eine vorliegende Zertifizierung des Repositoriums und gibt damit z.B. Auskunft darüber, ob ein Dienst nach anerkannten Archivierungsstandards betrieben wird (s. auch Abschnitt 7.5).   |

Quelle: Eigene Darstellung

Sofern Forschende nicht durch eine institutionelle Policy an die Veröffentlichung der Daten über einen bestimmten Dienst gebunden sind, können die folgenden Kriterien bei der Auswahl eines Repositoriums einbezogen werden.<sup>4</sup>

*Inhaltliche Ausrichtung:* Insbesondere disziplinäre Repositorien haben häufig einen Sammelschwerpunkt. D.h., sie archivieren vorrangig Forschungsdaten eines bestimmten Typs und/oder mit einem bestimmten thematischen Fokus. So archiviert etwa das GESIS Datenarchiv vorwiegend quantitative Daten, während qualitative Daten bei Qualiservice Bremen archiviert werden können.

*Technische Einschränkungen:* Archive, die Daten langfristig archivieren und kuratieren, beschränken häufig die Dateiformate, in denen Daten eingereicht werden können. Wenn die zu erhebenden Daten im Projekt nicht direkt in diesen Formaten generiert oder später verlustfrei in diese überführt werden können, muss ggf. ein anderer Dienst gewählt werden. Es empfiehlt sich aber im Zweifel, zunächst Rücksprache mit dem entsprechenden Dienst zu halten.

*Auffindbarkeit der Daten für die primäre Zielgruppe:* In der Regel haben disziplinspezifische Angebote den Vorteil, dass sie für die Fachcommunity relevante Daten bündeln und so eher Anlaufpunkte für die gezielte Suche nach Daten sind. Damit erreichen Forschungsdaten in diesen Angeboten die intendierte Zielgruppe in höherem Maße. Jener Nachteil institutioneller oder generischer Angebote kann aber ggf. abgeschwächt werden, wenn die beschreibenden Metadaten von anderen, auch disziplinspezifischen Angeboten, geharvested werden und gezielt durchsucht werden können.

*Zugangsbedingungen und -lizenzen:* Repositorien bieten teils unterschiedliche Zugangsbedingungen und Lizenzoptionen für Forschungsdaten. Diese müssen zu den Daten selbst sowie zu den Bedürfnissen der Datenproduzentinnen und -produzenten sowie der Nutzenden passen. So ermöglichen nicht alle Repositorien, Forschungsdaten für einen begrenzten Zeitraum unter ein Embargo zu setzen oder Zugangsbeschränkungen für sensitive Daten einzurichten. Beispielsweise erlaubt Dryad eine Veröffentlichung von Daten nur unter einer CC0-Lizenz, d.h. ohne jegliche Einschränkungen der Nutzung (vgl. Dryad 2018). Haben For-

4 Eine Checkliste zur Auswahl eines geeigneten Repositoriums bietet Whyte (2015). Ein kriteriengeleiteter Vergleich ausgewählter Dienste findet sich bei Gautier und Murphy (2018).

schende Befragten zugesichert, dass ihre Daten nur wissenschaftlich genutzt werden dürfen, können diese Daten nicht unter einer CC0-Lizenz veröffentlicht werden.

*Kuratierungslevel:* Wie oben dargestellt, unterscheiden sich Angebote zur Sicherung und Bereitstellung von Daten häufig hinsichtlich der durchgeführten Kuratierung. Viele Angebote sichern und veröffentlichen Daten nur unverändert und mit den von den Datengebenden vergebenen Metadaten. Andere – unter Umständen auch kostenpflichtige – Dienste prüfen Qualität und Konsistenz der Daten, transformieren sie in Archivierungs- und Nutzungsformate und schaffen Mehrwerte, indem sie Daten umfassend erschließen und mit weiteren Informationen anreichern. Die Frage, welches Kuratierungslevel für die jeweiligen Daten angemessen ist, ist nicht immer einfach zu beantworten. Sie hängt einerseits von der Bedeutung und Einmaligkeit der Daten sowie zukünftigen Nutzungsszenarien, andererseits von den verfügbaren Mitteln ab. Hinweise zur Auswahl von Daten für die Archivierung gibt das Digital Curation Centre (2014).

*Services für Datengebende:* Existierende Dienste unterscheiden sich z.B. im Umfang der Beratungs- und Unterstützungsleistungen für Forschende. Ein weiteres Kriterium kann die Verfügbarkeit von Zugriffsstatistiken oder sogenannter Altmetrics sein (vgl. Neylon 2014), die für Forschende eine wichtige Rolle dabei spielen können, ihren wissenschaftlichen Impact zu dokumentieren. Auch die Einbindung des ORCID-Diensts kann ein für Forschende relevanter Mehrwert sein.

*Kosten:* Es gibt eine große Bandbreite an kostenlosen Diensten, sowohl von öffentlich finanzierten Einrichtungen als auch von kommerziellen Anbietern. Kosten für die Datengebenden können z.B. dann entstehen, wenn seitens des Repositoriums die Daten noch aufwändig bearbeitet oder mit Informationen angereichert werden oder wenn ein großer Speicherbedarf besteht. Gleichfalls können Kosten für besondere Formen der Zugänglichmachung anfallen. Solche Kosten sollten idealerweise direkt bei der Projektplanung berücksichtigt werden und in Anträge auf Projektförderung Eingang finden.

*Qualität und Reputation des Diensts:* Um zu einer Einschätzung der Qualität und Reputation zu gelangen, können Forschende auf folgende Aspekte achten: Wird das Repositorium von Fachgesellschaften, Zeitschriften, Verlagen empfohlen? Zudem können bei der Beantwortung der Frage, ob ein Dienst in der Lage ist, die langfristige Sicherung und Kuratierung mit dem Ziel des Erhalts der Nutzbarkeit der Forschungsdaten zu leisten, vorliegende Zertifizierungen und Akkreditierungen hilfreich sein. Es existiert eine Reihe von Verfahren, die die Qualität und fachliche Expertise von Repositorien bewerten. Das Data Seal of Approval (DSA) und ICSU World Data System (WDS), welche unterdessen vom CoreTrustSeal (CTS) abgelöst wurden, sowie das (nicht forschungsdatenspezifische) nestor-Siegel für vertrauenswürdige digitale Langzeitarchive setzen einen Schwerpunkt beim Aspekt der Langzeitarchivierung, beziehen aber dennoch disziplinspezifische Aspekte und die Dimension der Nutzbarkeit archivierter Objekte mit ein. Das Akkreditierungsverfahren des Rats für Sozial- und Wirtschaftsdaten (RatSWD) hingegen legt einen starken Fokus auf fachspezifische Aspekte sowie auf die Bereitstellung von Forschungsdaten für die wissenschaftliche Nutzung. Erwähnt werden soll hier auch das DINI-Zertifikat. Dieses richtet sich hauptsächlich an sogenannte Hochschulschriftenserver, d.h. primär Repositorien für Textpublikationen. Es evaluiert insbesondere Aspekte, die im Zusammenhang mit Open Access zu Publikationen von Bedeutung sind (Zugänglichkeit, Sichtbarkeit, etc.). Durch seine Ausrichtung hat das DINI-Zertifikat eine eher geringe Aussagekraft bezüglich der Qualität von Forschungsdatenrepositorien – dennoch kann es Hinweise darauf geben, ob ein Dienst seine Arbeitsabläufe kritisch reflektiert hat und Mechanismen zur Qualitätssicherung bestehen.

Sofern möglich, sollten Forschende bei der Auswahl eines geeigneten Dienstes fachlichen Repositorien den Vorzug geben, die sich einer der erwähnten Zertifizierungen oder Akkreditierungen unterzogen haben. Damit können sie sicher sein, dass die entsprechenden Dienste

grundsätzlich in der Lage sind, Forschungsdaten fachgerecht und nachhaltig zu kuratieren und für die wissenschaftliche Nutzung bereitzustellen. Zudem erfüllen zertifizierte Dienste in der Regel die Anforderungen von Forschungsförderern, die zunehmend verlangen, dass Forschungsdaten nach Abschluss der Förderphase zur Verfügung gestellt werden.

## 7.6 Fazit

Der Fokus von Forschenden ist in erster Linie darauf gerichtet, die für die jeweiligen Forschungsfragen benötigten Daten zu erzeugen, zu verwalten und auszuwerten. Um die Nachvollziehbarkeit der Forschungsergebnisse zu gewährleisten und/oder die generierten Forschungsdaten Dritten für die Nutzung in neuen Kontexten und Projekten zugänglich zu machen, muss aber auch die Zeit nach der aktiven Projektphase in der Forschungsdatenmanagementplanung berücksichtigt und gezielt vorbereitet werden. Hierzu sollten Forschende sich frühzeitig eine Orientierung über die vielfältigen, teils neuen Möglichkeiten verschaffen, ihre Daten (langfristig) zu sichern und zu veröffentlichen und insbesondere die reichhaltigen Informations- und Beratungsangebote in Anspruch nehmen.

## Literaturverzeichnis

- ALLEA – All European Academies (2017): The European Code of Conduct for Research Integrity. Revised Edition. [https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020-ethics\\_code-of-conduct\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020-ethics_code-of-conduct_en.pdf) [Zugriff: 17.04.2018].
- Allianz der deutschen Wissenschaftsorganisationen (2010): Grundsätze zum Umgang mit Forschungsdaten. <http://www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/grundsaeetze/> [Zugriff: 17.04.2018].
- BMBF – Bundesministerium für Bildung und Forschung (2017): Bekanntmachung. Richtlinie zur Förderung von Forschung zu Digitalisierung im Bildungsbereich – Grundsatzfragen und Gelingensbedingungen. In: Bundesanzeiger, AT. 26.09.2017.
- Creative Commons – Mehr über die Lizenzen (o.J.). <https://creativecommons.org/licenses/> [Zugriff: 24.08.2018].
- DFG – Deutsche Forschungsgemeinschaft (2015): Umgang mit Forschungsdaten. DFG-Leitlinien zum Umgang mit Forschungsdaten. [http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien\\_forschungsdaten.pdf](http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf) [Zugriff: 13.04.2018].
- DFG – Deutsche Forschungsgemeinschaft (2013): Sicherung guter wissenschaftlicher Praxis. Denkschrift. Weinheim: Wiley-VCH. [http://www.dfg.de/download/pdf/dfg\\_im\\_profil/reden\\_stellungnahmen/download/empfehlung\\_wiss\\_praxis\\_1310.pdf](http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf) [Zugriff: 13.04.2018].
- Digital Curation Centre (o.J.): What is Digital Curation? <http://www.dcc.ac.uk/digital-curation/what-digital-curation> [Zugriff: 13.04.2018].
- Digital Curation Centre (2014): Five Steps to Decide What Data to Keep. A Checklist for Appraising Research Data. Volume 1. <http://www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep> [Zugriff: 13.04.2018].
- Dryad (2018): Frequently Asked Questions. What Kinds of Data does Dryad Accept? <https://datadryad.org/pages/faq> [Zugriff: 13.04.2018].
- Gautier, Julian/Murphy, Derek (2018): Comparative Review of Data Repositories. <https://doi.org/10.7910/DVN/WS9OUR>.
- Herb, Ulrich (2015): Open Science in der Soziologie. Eine interdisziplinäre Bestandsaufnahme zur offenen Wissenschaft und eine Untersuchung ihrer Verbreitung in der Soziologie. Glückstadt: Werner Hülsbusch. <https://doi.org/10.5281/zenodo.31234>.
- ISO – International Standards Organisation (2012): ISO 14721:2012 (CCSDS 650.0-P-1.1). Space Data and Information Transfer Systems. Open Archival Information System (OAIS). Reference model.

- Kindling, Maxi/Schirmbacher, Peter/Simukovic, Elena (2013): Forschungsdatenmanagement an Hochschulen. Das Beispiel der Humboldt-Universität zu Berlin. In: LIBREAS. Library Ideas 23, S. 43-63. <https://doi.org/10.18452/9041>.
- Kreutzer, Till (2016): Open Content. Ein Praxisleitfaden zur Nutzung von Creative-Commons-Lizenzen. <https://irights.info/artikel/neue-version-open-content-ein-praxisleitfaden-zu-creative-commons-lizenzen/26086> [Zugriff: 13.04.2018].
- Kuipers, Tom/Hoeven van der, Jeffrey (2009): Insight into digital preservation of research output in Europe. Survey Report. PARSE.Insight D3.4. <http://libereurope.eu/wp-content/uploads/2010/01/PARSE.Insight.-Deliverable-D3.4-Survey-Report.-of-research-output-Europe-Title-of-Deliverable-Survey-Report.pdf> [Zugriff: 17.04.2018].
- Leibniz-Gemeinschaft (2015): Empfehlungen der Leibniz-Gemeinschaft zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit Vorwürfen wissenschaftlichen Fehlverhaltens. <https://www.leibniz-gemeinschaft.de/forschung/gute-wissenschaftliche-praxis/> [Zugriff: 17.04.2018].
- Mauer, Reiner (2012): Das GESIS Datenarchiv für Sozialwissenschaften. In: Althenhöner, Reinhard/Oellers, Claudia (Hrsg.): Langzeitarchivierung von Forschungsdaten. Standards und disziplinspezifische Lösungen. Berlin: Scivero, S. 197-215. [https://www.ratswd.de/dl/downloads/langzeitarchivierung\\_von\\_forschungsdaten.pdf](https://www.ratswd.de/dl/downloads/langzeitarchivierung_von_forschungsdaten.pdf) [Zugriff: 17.04.2018].
- Mochmann, Ekkehard (2008): Improving the Evidence Base for International Comparative Research. In: International Social Science Journal 59, 193/194, S. 489-506.
- NISO – National Information Standards Organization (2013): Recommended Practices for Online Supplemental Journal Article Materials. [https://groups.niso.org/apps/group\\_public/download.php/10055/RP-15-2013\\_Supplemental\\_Materials.pdf](https://groups.niso.org/apps/group_public/download.php/10055/RP-15-2013_Supplemental_Materials.pdf) [Zugriff: 17.04.2018].
- nestor-Arbeitsgruppe OAIS-Übersetzung/Terminologie (2013): Referenzmodell für ein Offenes Archiv-Informationssystem. Deutsche Übersetzung 2.0. nestor Materialien 16. <http://nbn-resolving.de/urn:nbn:de:0008-2013082706> [Zugriff: 17.04.2018].
- Neylon, Cameron (2014): Altmetrics: What are they good for? <http://blogs.plos.org/opens/2014/10/03/altmetrics-what-are-they-good-for/> [Zugriff: 13.04.2018].
- Open Data Institute (o.J.): The Data Spectrum. <https://theodi.org/about-the-odi/the-data-spectrum/> [Zugriff: 13.04.2018].
- Pennock, Maureen (2006): JISC Briefing Paper. Digital Preservation. Continued Access to Authentic Digital Assets. [https://www.webarchive.org.uk/wayback/archive/20140614202005/http://www.jisc.ac.uk/publications/briefingpapers/2006/pub\\_digipreservationbp.aspx#main-content-area](https://www.webarchive.org.uk/wayback/archive/20140614202005/http://www.jisc.ac.uk/publications/briefingpapers/2006/pub_digipreservationbp.aspx#main-content-area) [Zugriff: 17.04.2018].
- Piwowar, Heather/Vision, Todd J. (2013): Data Reuse and the Open Data Citation Advantage. In: PeerJ 1:e175. <https://doi.org/10.7717/peerj.175>.
- Quandt, Markus/Mauer, Reiner (2012): Sozialwissenschaften. In: Neuroth, Heike/Strathmann, Stefan/Oßwald, Achim/Scheffel, Regine/Klump, Jens/Ludwig, Jens (Hrsg.): Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme. Boizenburg: Werner Hülsbusch, S. 61-81. [http://www.nestor.sub.uni-goettingen.de/bestandsaufnahme/nestor\\_lza\\_forschungsdaten\\_bestandsaufnahme.pdf](http://www.nestor.sub.uni-goettingen.de/bestandsaufnahme/nestor_lza_forschungsdaten_bestandsaufnahme.pdf) [Zugriff: 17.04.2018].
- Recker, Jonas/Zenk-Möltgen, Wolfgang/Mauer, Reiner (2017): 7. Applications of Research Data Management at GESIS Data Archive for the Social Sciences. In: Kruse, Filip/Thestrup, Jesper Boserup (Hrsg.): Research Data Management - A European Perspective. Berlin, Boston: De Gruyter, S. 119-146.
- Reilly, Susan/Schallier, Wouter/Schrimpf, Sabine/Smit, Eefke/Wilkinson, Max (2011): ODE – Opportunities for Data Exchange. Report on Integration of Data and Publications. <https://doi.org/10.5281/zenodo.8307>.
- Rücknagel, Jessika/Vierkant, Paul/Ulrich, Robert/Kloska, Gabriele/Schnepf, Edeltraut/Fichtmüller, David/Reuter, Evelyn/Semrau, Angelika/Kindling, Maxi/Pampel, Heinz/Witt, Michael/Fritze, Florian/van de Sandt, Stephanie/Klump, Jens/Goebelbecker, Hans-Jürgen/Skarupianski, Michael/Bertelmann, Roland/Schirmbacher, Peter/Scholze, Frank/Kramer, Claudia/Fuchs, Claudio/Spier, Shaked/Kirchhoff, Agnes (2015): Metadata Schema for the Description of Research Data Repositories. Version 3.0. <https://doi.org/10.2312/re3.008>.
- Scheuch, Erwin K. (2003): History and Visions in the Development of Data Services for the Social Sciences. In: International Social Science Journal 55, 177, S. 385-399.
- Schiller, David H./Eberle, Johanna/Fuß, Daniel/Goebel, Jan/Heining, Jörg/Mika, Tatjana/Müller, Dana/Röder, Frank/Stegmann, Michael/Stephan, Karsten (2017): Standards des sicheren Datenzugangs in den Sozial- und Wirtschaftswissenschaften. Überblick über verschiedene Remote-Access-Verfahren. RatSWD Working Paper Series 261. [https://www.ratswd.de/dl/RatSWD\\_WP\\_261.pdf](https://www.ratswd.de/dl/RatSWD_WP_261.pdf) [Zugriff: 17.04.2018].
- Steinhau, Henry/Pachali, David (2017): Was ist Creative Commons Zero? <https://irights.info/artikel/was-ist-cc0/28750> [Zugriff: 17.04.2018].

Whyte, Angus (2015): Where to Keep Research Data. DCC Checklist for Evaluating Data Repositories v.1.1. <http://www.dcc.ac.uk/resources/how-guides-checklists/where-keep-research-data/where-keep-research-data> [Zugriff: 17.04.2018].

## Linkverzeichnis

CESSDA – Consortium of European Social Science Data Archives: <https://www.cessda.eu/> [Zugriff: 16.04.2018].

CTS – CoreTrustSeal: <https://www.coretrustseal.org/> [Zugriff: 16.04.2018].

Cryptshare: <https://www.cryptshare.com/de/start/> [Zugriff: 04.2018].

Data Documentation Initiative: <https://www.ddalliance.org/> [Zugriff: 16.04.2018].

DSA – Data Seal of Approval: <https://www.datasealofapproval.org/en/> [Zugriff: 16.04.2018].

DBK – Datenbestandskatalog: <https://dbk.gesis.org/dbksearch/home.asp?db=d> [Zugriff: 16.04.2018].

DINI-Zertifikat: <https://dini.de/dini-zertifikat/> [Zugriff: 16.04.2018].

Dryad: <https://datadryad.org/> [Zugriff: 05.06.2018].

ESSD – Earth System Science Data: Interactive Peer Review: [https://www.earth-system-science-data.net/peer\\_review/interactive\\_review\\_process.html](https://www.earth-system-science-data.net/peer_review/interactive_review_process.html) [Zugriff: 16.04.2018].

Figshare: <https://figshare.com/> [Zugriff: 16.04.2018].

Forschungsdaten.org: Datenjournale: [http://www.forschungsdaten.org/index.php/Data\\_Journals](http://www.forschungsdaten.org/index.php/Data_Journals) [Zugriff: 16.04.2018].

Forschungsdaten.org: Policies: [http://www.forschungsdaten.org/index.php/Data\\_Policies#Institutionelle\\_Policies](http://www.forschungsdaten.org/index.php/Data_Policies#Institutionelle_Policies) [Zugriff: 16.04.2018].

GESIS – Datenarchiv für Sozialwissenschaften: <http://www.gesis.org/institut/abteilungen/datenarchiv-fuer-sozialwissenschaften/> [Zugriff: 05.06.2018].

SDC – GESIS – Secure Data Center: <https://www.gesis.org/sdc> [Zugriff: 16.04.2018].

WDS – ICSU-World Data System: <https://www.icsu-wds.org/services/certification> [Zugriff: 16.04.2018].

Leibniz-Zentrum für Psychologische Information und Dokumentation – PsychData: <https://leibniz-psychology.org/angebote/archivieren/> [Zugriff: 05.06.2018].

Mendeley Data: <https://data.mendeley.com/> [Zugriff: 16.04.2018].

Nature – Recommended Data Repositories: <https://www.nature.com/sdata/policies/repositories> [Zugriff: 16.04.2018].

nestor-Siegel für vertrauenswürdige digitale Langzeitarchive: [http://www.langzeitarchivierung.de/Subsites/nestor/DE/Siegel/siegel\\_node.html](http://www.langzeitarchivierung.de/Subsites/nestor/DE/Siegel/siegel_node.html) [Zugriff: 16.04.2018].

ORCID - Open Researcher and Contributor ID: <https://orcid.org/> [Zugriff: 16.04.2018].

PLOS – Recommended Repositories: <http://journals.plos.org/plosone/s/data-availability#loc-recommended-repositories> [Zugriff: 16.04.2018].

Qualiservice Bremen: <http://www.qualiservice.org> [Zugriff: 05.06.2018].

RADAR: <https://www.radar-service.eu/> [Zugriff: 05.06.2018].

RatSWD – Rat für Sozial- und Wirtschaftsdaten: <https://www.ratswd.de> [Zugriff: 16.04.2018].

RatSWD: Akkreditierungsverfahren: <https://www.ratswd.de/forschungsdaten/qualitaetsicherung> [Zugriff: 16.04.2018].

RatSWD: Forschungsdatenzentrum: <https://www.ratswd.de/forschungsdaten/FDZ> [Zugriff: 16.04.2018].

Re3Data: <https://www.re3data.org/> [Zugriff: 16.04.2018].

Research Data Journal for the Humanities and Social Sciences: <http://www.brill.com/products/online-resources/research-data-journal-humanities-and-social-sciences> [Zugriff: 16.04.2018].

RISources: [http://risources.dfg.de/home\\_de.html](http://risources.dfg.de/home_de.html) [Zugriff: 16.04.2018].

Robert Koch-Institut, Forschungsdatenzentrum: [https://www.rki.de/DE/Content/Gesundheitsmonitoring/Forschungsdatenzentrum/forschungsdatenzentrum\\_node.html](https://www.rki.de/DE/Content/Gesundheitsmonitoring/Forschungsdatenzentrum/forschungsdatenzentrum_node.html) [Zugriff: 05.06.2018].

Universität Bielefeld – Datenservicezentrum Betriebs- und Organisationsdaten: <http://www.uni-bielefeld.de/dsz-bo/> [Zugriff: 05.06.2018].

Universität Heidelberg – heiDATA: <https://heidata.uni-heidelberg.de/> [Zugriff: 05.06.2018].

Zenodo: <https://zenodo.org> [Zugriff: 05.06.2018].

