
eDissPlus – Optionen für die Langzeitarchivierung dissertationsbezogener Forschungsdaten aus Sicht von Bibliotheken und Forschenden

Dirk Weisbrod¹, Ben Kaden², Michael Kleineberg³

1 Deutsche Nationalbibliothek

2,3 Universitätsbibliothek, Humboldt-Universität zu Berlin

Zusammenfassung. Im Rahmen des DFG-Projektes “Elektronische Dissertationen Plus” (eDissPlus) beschäftigen sich die Universitätsbibliothek der Humboldt-Universität zu Berlin (HU) und die Deutsche Nationalbibliothek (DNB) mit konzeptionellen und technischen Anforderungen für eine zeitgemäße Archivierung und Publikation von Forschungsdaten, die im Zusammenhang mit Promotionsvorhaben entstehen. Der Beitrag diskutiert zunächst die Verantwortlichkeiten für die Veröffentlichung und Langzeitarchivierung dissertationsbezogener Forschungsdaten auf nationaler Ebene und über Community-Grenzen hinweg. Dabei stellen sich insbesondere die Fragen, welche Forschungsdaten überhaupt zu einer Dissertation gehören und unter welchen Umständen die DNB für deren Langzeitarchivierung verantwortlich sein sollte. Anschließend werden Zwischenergebnisse der Anforderungsanalyse vorgestellt, die auf Interviews mit Promovierenden, Post-Docs und Gutachtern aus unterschiedlichen Wissenschaftsbereichen an der Humboldt-Universität zu Berlin basieren.

Schlagwörter. Elektronische Dissertationen, Forschungsdaten, Langzeitarchivierung, Forschungsdatenpublikation, Bibliotheken

Einleitung

Entsprechend der sehr ausdifferenzierten fachspezifischen Datenerhebung bzw. -verarbeitung erfolgt die Veröffentlichung und Archivierung von Forschungsdaten in der Regel in Community- oder fachspezifischen Repositorien und nicht Community-übergreifend.¹ Es gibt aber auch Ausnahmen. So ergibt sich aus den gesetzlichen Vorgaben der Deutschen Nationalbibliothek (DNB) der Auftrag, Forschungsdaten, die Teil einer Publikation sind, auf nationaler Ebene zu sammeln. Eine Lösung für diesen Auftrag erarbeitet die DNB derzeit zusammen mit der Humboldt-Universität zu Berlin (HU) im DFG-Projekt „Elektronische Dissertationen Plus“ (eDissPlus), das sich auf einen Teilbereich dieses Komplexes konzentriert, nämlich auf Forschungsdaten, die Promovierende im Rahmen ihres Dissertationsprojekts generieren und veröffentlichen. Damit erfasst eDissPlus die ganze Bandbreite an Disziplinen und Forschungsprojekten, die an einer promotionsberechtigten Hochschule existieren, sodass eine Adaption der Projektergebnisse auf andere

1 So verzeichnet das Verzeichnis re3data.org (Stand Februar 2017) hauptsächlich disziplinspezifische Einrichtungen (1560). Die anderen recherchierbaren Angaben sind institutionell (465) und „other“ (189). Angaben zu Community- oder fachübergreifenden Repositorien liegen nicht vor. Verfügbar unter: <http://www.re3data.org/search>.

Community-übergreifende Szenarien möglich ist. Unter anderem sollen in diesem Projekt ein Ingest-Prozess für die Publikation von elektronischen Dissertationen mit Forschungsdaten entwickelt und der Pflichtablieferungsworkflow der DNB in Hinblick auf die neuen Sammelobjekte erweitert werden. Der vorliegende Beitrag präsentiert Erkenntnisse, die in den ersten Monaten des Projekts gewonnen wurden und für die Forschungsdaten-Community von Interesse sind. Es handelt sich erstens um Überlegungen zu den Verantwortungsstrukturen, die aus Sicht der DNB bei der Langzeitarchivierung publikationsbezogener Forschungsdaten zum Einsatz kommen sollten und zweitens um die Ergebnisse von Interviews, die für eine Anforderungsanalyse auf Seiten der HU vor allem mit Promovierenden geführt wurden.

Verantwortungsstrukturen

Die DNB und ihr Sammelauftrag

Der Sammelauftrag der Deutschen Nationalbibliothek ist in der Verordnung über die Pflichtablieferung von Medienwerken an die Deutsche Nationalbibliothek (Pflichtablieferungsverordnung - PflAV) definiert. Diese enthält auch klare Aussagen über Forschungsdaten. So definiert § 9 (Weitere Einschränkungen der Ablieferungspflicht für Netzpublikationen), dass selbständig veröffentlichte Primär-, Forschungs- und Rohdaten nicht abzuliefern sind. Das impliziert eine Ablieferungspflicht für unselbständig publizierte Forschungsdaten, die laut § 7 abzuliefern sind, da „die Ablieferungspflicht auch alle Elemente, Software und Werkzeuge [umfasst], die in physischer oder in elektronischer Form erkennbar zu den ablieferungspflichtigen Netzpublikationen gehören, auch wenn sie für sich allein nicht der Ablieferungspflicht unterliegen“ (BMJV 2017).

Damit wird deutlich, dass die Zuständigkeit der DNB sich nicht auf alle Publikationstypen von Forschungsdaten erstreckt. Vielmehr ist ihr Sammelauftrag auf Daten beschränkt, die „erkennbar“ zu einer Netzpublikation gehören und damit unselbständig publiziert wurden. Aus diesen Vorgaben wurden in der ersten Projektphase Kriterien für eine noch zu schaffende Forschungsdaten-Policy der DNB abgeleitet. Es galt zu konkretisieren, wann Forschungsdaten erkennbar zu einer Dissertation gehören und in welchen Fällen die DNB oder andere Institutionen für deren Langzeitarchivierung verantwortlich sind.

Die DNB konnte sich nicht bereits bestehende Verantwortungsstrukturen in andern Ländern zum Vorbild nehmen, da vergleichbare Projekte offenbar noch nicht durchgeführt wurden.² Eine Ausnahme bilden Projekte, die sich mit der Verknüpfung von wissenschaftlicher Publikation, Daten und Autoreninformationen mittels persistenter Identifizierung (PID) beschäftigen. Ein Beispiel hierfür ist THOR (Fenner et al. 2015). In Baden-Württemberg gibt es zudem das Projekt bwDataDiss, das allerdings andere Schwerpunkte verfolgt (ALWR 2017).

2 Einen Überblick über die Aktivitäten europäischer Länder gibt die Publikation der Europäischen Kommission: „Access to and Preservation of Scientific Information in Europe“ (Tarazona Rúa et al. 2015). Sie verdeutlicht, dass je nach Nation verschiedene Akteure für das Sammeln und den Zugang zu wissenschaftlichen Publikationen zuständig sind, was den Vergleich mit den deutschen Gegebenheiten sehr erschwert.

Welche Forschungsdaten gehören zu einer Dissertation?

Die „erkennbare Zugehörigkeit“ kann durch den Integrationsgrad der Daten in die Dissertation bestimmt werden. Ein Blick auf die Data Publication Pyramid, die im Rahmen des EU-Projektes Opportunities für Data Exchange (ODE) unter Teilnahme der DNB erarbeitet wurde (DNB 2013), visualisiert die möglichen Integrationsgrade (Reilly et al. 2011, 5). Die Integration von Daten und Publikation nimmt demnach von oben nach unten ab, bis im unteren Teil der Pyramide reine Datenpublikationen als selbständig publizierte Daten (Typ 4) sowie unpublizierte Daten (Typ 5) stehen, bei denen kein Konnex zu einer separaten Publikation existiert und die daher nicht unter den Sammelauftrag der DNB fallen. Interessant für den Sammelauftrag der DNB sind hingegen die Typen 1 bis 3.

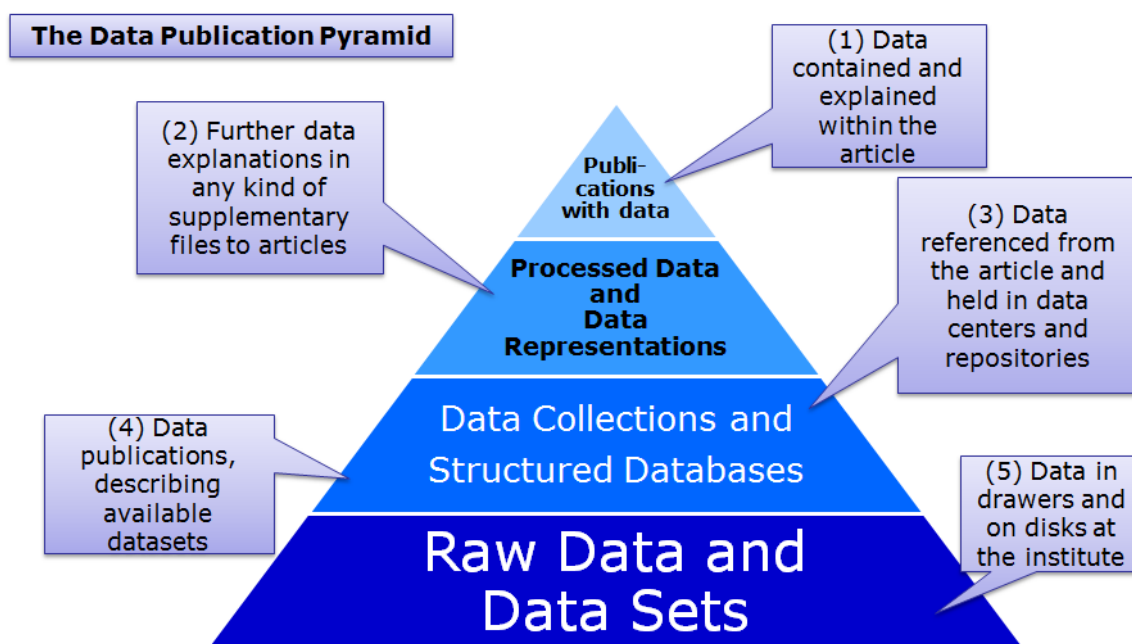


Abbildung 1. The Data Publication Pyramid (Reilly et al. 2011, 6).

Die Spitze der Pyramide bilden vollständig in eine Dissertationsschrift integrierte Daten (Typ 1). Das sind zum Beispiel Grafiken oder Tabellen, die in ein PDF eingebettet oder einer Printversion abgedruckt wurden. Forschungsdaten, die in dieser Form vorliegen, sammelte die DNB schon immer – sozusagen automatisch als Bestandteil der abgabepflichtigen Hochschulschrift. Auf diesen Typ musste eDissPlus deshalb auch nicht näher eingehen. Typ 2 beschreibt Daten, die als Supplemente einer Dissertation beigelegt sind. Varianten dieses Typs sind gedruckte Dissertationen, denen CD-ROMs oder andere Datenträger mit Forschungsdaten beiliegen. Überträgt man dieses Beispiel auf Netzpublikationen, dann manifestiert sich die erkennbare Zugehörigkeit der Forschungsdaten zur Dissertation darin, dass sie als Supplemente von Promovierenden zusammen mit der Dissertation auf den Publikationsserver der Hochschule geladen wurden. Der Integrationsgrad der Daten in die Dissertation ist immer noch hoch. Es ist anzunehmen, dass die beigegebenen Daten in enger Verbindung zur Dissertation und den darin präsentierten Forschungsergebnissen stehen. Supplemente sollten daher auch von der DNB gesammelt werden. Typ 3 illustriert die Möglichkeit, dass Forschungsdaten auf einem externen Repositorium liegen, in der Dissertati-

on aber referenziert werden (etwa durch Angabe einer PID). Man kann nun fragen, ob es sich hierbei noch um Daten handelt, die erkennbar zu der betreffenden Dissertation gehören und somit unter den Sammelauftrag der DNB fallen. Denn diese Daten gelten, sofern Sie über eine PID verfügen, zugleich auch als selbständig publizierte Daten. Deshalb werden die Typen 3 und 4 in der Data Publication Pyramid auch innerhalb einer Schicht verortet.

Umgang mit referenzierten Daten

Eine Lösung für den Umgang mit referenzierten Daten eröffnet der inhaltliche Bezug von Daten und Dissertation. Dieser manifestiert sich in dem Beitrag der Forschungsdaten zum Verständnis der Dissertation, etwa indem sie Argumentationen oder Hypothesen der Promovierenden stützen oder falsifizieren und bestimmte Sachverhalte illustrieren. Eng damit verbunden ist der Anspruch, mithilfe der vorhandenen Daten die präsentierten Forschungsergebnisse reproduzieren und damit den Forschungsprozess selbst verstehen und nachvollziehen zu können. Ein denkbare Anwendungsfeld ist die Aufdeckung wissenschaftlichen Fehlverhaltens, so wie es die Empfehlungen der DFG zur Sicherung guter wissenschaftlicher Praxis beschreiben (DFG 2013). Folgt man dieser Argumentation, dann ist auch die Sichtweise berechtigt, dass referenzierte Daten ebenso wie Supplemente in einem engen inhaltlichen Bezug zur Dissertation stehen können. Allerdings entsteht dieser Bezug nicht durch eine einfache Datenzitation im Anmerkungsapparat des Dissertationstextes. Es könnte sich dabei auch um Daten handeln, die einem vergleichbaren Forschungsdesign entstammen, aber für die Dissertation nicht ausgewertet wurden. Um die erkennbare Zugehörigkeit der externen Daten zu einer Dissertation zu signalisieren, sollten die Promovierenden die Referenzen während des Abgabeprozesses der Dissertation explizit als solche ausweisen. Das kann durch den Eintrag von PIDs (z. B. DOI, URN, handle) in ein Webformular geschehen. In einem solchen Fall können die extern archivierte Daten als erkennbarer Bestandteil der Dissertation betrachtet und von der DNB gesammelt oder doch zumindest im Metadatensatz der Dissertation durch Angabe der PID referenziert werden.

Damit wird eines deutlich: Letztlich entscheiden die Promovierenden über die Bedeutung der Daten und ihre Unverzichtbarkeit im Hinblick auf Verständnis und Nachvollziehbarkeit der in der Dissertation publizierten Forschungsergebnisse.³ Daher ist es Aufgabe der abliefernden Hochschulbibliotheken, durch propädeutische Maßnahmen (z. B. Schulungen, Workshops, Guidelines) bei den Studierenden ein Bewusstsein für die Bedeutung von Forschungsdaten zu erzeugen. Die Entwicklung entsprechender Maßnahmen ist folglich auch Teil des Projektes eDissPlus.

Zudem muss die Frage geklärt werden, ob die DNB extern archivierte Daten auch in ihr Langzeitarchiv übernimmt oder ob ein Metadatenachweis im Katalog ausreicht, zumal Forschungsdaten bereits vor Ablieferung der Dissertationsschrift auf einem Repositorium vorliegen können.⁴ Die im vergangenen Jahr erschienenen Empfehlungen des Rates für Informationsinfrastrukturen (RfII) weisen hier eindeutig die Richtung, indem sie einen verteilten Archivierungsansatz vorschlagen, wonach „Aufgaben der Langzeitarchivierung als informationsinfrastrukturelle Daueraufgaben durch ggf. miteinander vernetzte Einrichtungen im Rahmen der Nationalen For-

3 Ohnehin tragen die Datenproduzenten, in diesem Fall die Promovierenden, Verantwortung für die Datenqualität, wozu im Falle eines Dissertationsprojektes auch die Auswahl der beizufügenden Daten gehören sollte. Vgl. hierzu auch die Empfehlungen des Rates für Informationsinfrastrukturen (RfII 2016).

4 Vergleiche hierzu auch die Punkte 6 und 10 der unten präsentierten Interview-Zwischenergebnisse.

schungsdateninfrastruktur (NFDI) zu leisten und zu koordinieren“ sind (RfII 2016, 47). Externe Daten sollten somit grundsätzlich nicht von der DNB übernommen, sondern in den bereits genutzten Repositorien verbleiben und lediglich im Metadatensatz der betreffenden Dissertation referenziert werden. Allerdings ist zu beachten, dass Langzeitarchivierung oft missverständlich mit den Empfehlungen der Deutschen Forschungsgemeinschaft (DFG) gleichgesetzt wird, wonach Primärdaten zehn Jahre aufbewahrt werden sollen (DFG 2013). Langzeitarchivierung meint jedoch wesentlich längere Zeiträume und kennt auch kein Ablaufdatum (Neuroth et al. 2009). Es wird somit seitens der DNB zu prüfen sein, ob ein Repository im Sinne der Langzeitarchivierung vertrauenswürdig ist und über eine entsprechende Zertifizierung verfügt (z. B. Data Seal of Approval; nestor-Siegel). Fehlt eine solche Zertifizierung, müsste die DNB entscheiden, ob und wie sie die Daten übernimmt und archiviert. Da dies jedoch aufwendige Prüfroutinen und Schnittstellenentwicklungen voraussetzt, ist eine Lösung dieser Aufgabe im Rahmen von eDissPlus nicht wahrscheinlich. Die Übernahme und Langzeitarchivierung von Supplementen sollte aufgrund des größeren Integrationsgrades von Dissertation und Daten Vorrang haben.

Anforderungsanalyse

Erhebungsmethoden und Analyseverfahren

Während die DNB vor allem die Perspektive einer Archivbibliothek einnimmt, verfolgt die Universitätsbibliothek der HU eine stärker forschungszentrierte Sichtweise auf den Umgang mit Forschungsdaten. Um dementsprechend die Anforderungen der Promovierenden aus verschiedenen Fachdisziplinen zu ermitteln, werden derzeit zusätzlich zur Auswertung bestehender quantitativer Studien (Simukovic et al. 2013; KIT 2015; Tenopir et al. 2017) eine Reihe von qualitativen Leitfadeninterviews mit Promovierenden, Post-Docs sowie Gutachtenden von Dissertationen an der Humboldt-Universität zu Berlin durchgeführt. Interessant sind dabei – neben dem allgemeinen Umgang mit dissertationsbegleitenden Daten während des Forschungsprozesses – vor allem die Möglichkeiten ihrer Veröffentlichung. In den Interviews werden die für das jeweilige Dissertationsprojekt relevanten Forschungsdaten unter Berücksichtigung der möglichen Datentypen und Bearbeitungsphasen erhoben. Weiterhin wird nach spezifischen Rahmenbedingungen, der persönlichen Motivation, den Besonderheiten der fachkulturellen Praxis, der Evaluation bestehender Infrastrukturangebote sowie nach konkreten Desideraten gefragt.

Für die Befragungen wurden potentielle Interviewpartner über diverse Kommunikationskanäle wie die Humboldt Graduate School, Fachreferenten und Bibliotheksbeauftragte kontaktiert. Trotz großer Bemühungen war das Echo erstaunlich gering, weshalb die Interviewphase noch nicht abgeschlossen werden konnte. Zum Stand dieses Beitrags lagen 16 transkribierte Interviews vor aus den Fachbereichen Bibliotheks- und Informationswissenschaft, Erziehungswissenschaft, Europäische Ethnologie, Geografie, Geschichtswissenschaft, Kultur- und Sozialanthropologie, Kulturwissenschaft, Literaturwissenschaft, Medizin, Physik, Rechtswissenschaft sowie Soziologie. Zudem ist die Repräsentativität insofern eingeschränkt, als dass sich bislang hauptsächlich Personen mit einem besonderen persönlichen Interesse am Thema zum Gespräch bereit erklärten.

Für die Datenauswertung wurden aus den Interviewprotokollen jeweils zwischen 50 bis 80 relevante Einzelaussagen extrahiert, so dass bislang etwa 1100 Statements als Datenbasis zur Ver-

fügung stehen. Anhand einer qualitativen Inhaltsanalyse mit der Vergabe von Keywords und unter Einsatz des Erschließungs- und Recherchewerkzeugs *Statement Finder*⁵ werden wiederkehrende Einstellungsmuster, Prozesse, Anforderungen und Desiderate geclustert sowie fachspezifische Besonderheiten identifiziert. Auf dieser Grundlage lassen sich typisierte Nutzungsszenarien und Anforderungsprofile modellieren, welche die unterschiedlichen Zielgruppen für die Archivierungs- bzw. Publikationsangebote der HU abbilden sollen.

Zwischenergebnisse

Zu den zentralen Erkenntnissen der Interviewphase gehört, dass gemessen am Ideal einer offenen Wissenschaft der transparente und nachhaltige Umgang mit Forschungsdaten erhebliche Herausforderungen zu bewältigen hat. Diese umfassen individuelle Aspekte (z.B. Motivation, Zeitressourcen, IT-Kompetenz), wissenschaftsstrukturelle (z.B. Konventionen, Promotionsordnungen, Anreizsysteme), infrastrukturelle (z.B. IT-Dienste, Publikationsworkflows) sowie rechtliche Herausforderungen (z.B. Datenschutz, Lizenzierung, Urheberrecht). Als erstes Zwischenergebnis lassen sich folgende Punkte festhalten:

1. *Thema Forschungsdaten randständig in Wissenschaftspraxis*

Eher unerwartet und daher sehr lehrreich ist die Einsicht, dass das Thema der dissertationsbegleitenden Forschungsdatenpublikation aus Sicht der Forschenden derzeit sehr randständig erscheint. Bereichsübergreifend gibt es für solche Veröffentlichungsformen keine gängige Praxis, sondern allenfalls persönlich motivierte Einzelfälle. Impulse durch die Begutachtenden gibt es ebenso selten wie Best-Practice-Beispiele. Rezeption oder Nachnutzung der Forschungsdaten sowie eine entsprechende Kreditierung durch die Fachgemeinschaften werden kaum erwartet. In Fachdiskursen wird das Thema Forschungsdatenpublikation vor allem im Zusammenhang mit Vorgaben von Förderinstitutionen (z.B. DFG, BMBF, ERC) wahrgenommen.

2. *Kaum erkennbare Rahmenbedingungen für Forschungsdatenmanagement*

Abgesehen von den externen Vorgaben der Förderinstitutionen sind selten fachspezifische Rahmenbedingungen oder gar konkrete Handlungsempfehlungen für den Umgang mit Forschungsdaten feststellbar. Insbesondere findet das Thema kaum Berücksichtigung in geltenden Promotionsordnungen. Unabhängig von der Frage einer Forschungsdatenpublikation wird zwar deutlich, dass auf die Einhaltung der 10-jährigen Aufbewahrungsfrist von Forschungsdaten im Sinne der guten wissenschaftlichen Praxis Wert gelegt wird. Allerdings werden hierbei meist ad-hoc-Lösungen wie das Vorhalten der Daten auf privaten Speichermedien verfolgt. Promovierenden sind die spezifischen Ansprüche an das Forschungsdatenmanagement wie beispielsweise die Erstellung eines Datenmanagementplanes kaum bekannt.

3. *Speicherbedarfe und Datenformate überschaubar*

In Übereinstimmung mit einer vorherigen Studie (Simukovic et al. 2013) erreicht das Datenvolumen eines Promotionsprojektes selten mehr als einen zweistelligen Giga-

5 Der Statement Finder wurde im Rahmen des DFG-Projektes „Future Publications in den Humanities“ (Fu-Push) entwickelt und als Open-Source-Anwendung zur Nachnutzung bereitgestellt. Verfügbar unter: <https://www2.hu-berlin.de/fupush/statement-finder/#/statements>.

- bytebereich. Bei den Datenformaten besteht zwar durchaus eine gewisse Heterogenität, aber in der Regel werden gängige fachspezifische Standardformate verwendet.
4. *Teilweise hohe Anzahl an einzelnen Forschungsdaten (z.B. Quellenmaterialien)*
Wie granular Forschungsdaten als Publikationsobjekte anfallen steht in Abhängigkeit zum jeweiligen Forschungsprojekt. Es kann sich um eine einzelne Tabelle handeln oder um 8000 Abbildungen. Hier zeichnet sich Beratungsbedarf ab beispielsweise hinsichtlich des Bündelns zu bestimmten logischen Publikationseinheiten. Werden hohe Auffindbarkeit und separate Nachnutzung angestrebt, besteht Interesse an einer möglichst granularen Aufschlüsselung der Daten, einschließlich einer jeweiligen PID.
 5. *Forschungsdatenmanagement höher priorisiert als Forschungsdatenpublikation*
Während die Veröffentlichung von Forschungsdaten aus unterschiedlichen Gründen weitgehend nachgeordnet betrachtet wird, misst man dem Forschungsdatenmanagement durchgehend hohe Bedeutung zu. Dazu zählen aus Sicht der Promovierenden vor allem die Speicherung und Versionierung der Daten, die Verwaltung der Zugriffsrechte sowie die Dokumentation. Die Bedeutung von Datenmanagementplänen wird anerkannt, auch wenn die meisten Befragten keine entsprechenden Erfahrungen vorweisen. Großes Interesse besteht hinsichtlich der zum Teil unbekanntenen universitären Infrastruktur- und Beratungsangebote zum Forschungsdatenmanagement. Für Promovierende erweist sich die Perspektive einer Langzeitarchivierung bzw. -verfügbarkeit ihrer Forschungsdaten durch die DNB teilweise als Motivation, eine Forschungsdatenpublikation in Erwägung zu ziehen.
 6. *Forschungsdatenmanagement in datengetriebenen Disziplinen meist intern*
Datenintensive Fachbereiche wie Naturwissenschaften, Medizin, Geografie oder Digital Humanities bieten oft institutsinterne IT-Infrastrukturen und Lösungen zum Forschungsdatenmanagement, die als funktionierend bewertet und im Vergleich zu Angeboten der Universitätsbibliothek oder dem Rechenzentrum bevorzugt werden.
 7. *Hoher Beratungsbedarf in Bereichen qualitativer empirischer Sozialforschung*
Auch wenn die Befragungen selbst nur einen sehr selektiven Einblick bieten, zeichnet sich ab, dass in Fachbereichen mit einem hohen Anteil an qualitativer Sozialforschung wie in der Kultur- und Sozialanthropologie, der Europäischen Ethnologie oder der Sozialgeschichte ein besonderer Bedarf an Beratungs- und Dienstleistungsangeboten besteht. Dies wurde unter anderem auch in einem gemeinsamen Workshop mit dem Fachinformationsdienst (FID) „Kultur- und Sozialanthropologie“ an der Universitätsbibliothek der HU Berlin herausgearbeitet. Gründe dafür werden vor allem gesehen in der Heterogenität der in diesen Fächern auftretenden Forschungsdaten (z.B. Feldtagebücher, Interviews, Filmaufnahmen, Digitalisate von Artefakten und Archivalien), der Sensibilität der oft personenbezogenen und nur schwer anonymisierbaren Daten sowie einer im Vergleich zu klassischen MINT-Fächern weniger ausgeprägten IT-Affinität. Zum einen besteht Unsicherheit in der Frage, inwieweit ethnologische bzw. historische Quellenmaterialien überhaupt als Forschungsdaten angesehen werden können; zum anderen erscheint die Relevanz solcher Quellen und daher auch die Motivation einer Veröffentlichung besonders hoch.
 8. *Forschungsdatenpublikation stark abhängig von individueller Motivation*
Das Interesse am Thema Forschungsdaten scheint vor allem persönlich motiviert zu sein. Generell lassen sich für eine Forschungsdatenpublikation drei Motivationstypen

unterscheiden: a) eine pragmatische, bei der Forschungsdaten aus situativer Notwendigkeit und eventuell zur Transparenzsicherung der Dissertationsschrift begleitend publiziert werden; b) eine reputationale, bei der die Forschungsdatenpublikation in der Erwartung zusätzlicher wissenschaftlicher Kreditierung veröffentlicht wird und schließlich c) eine ethische, die Zugänglichkeit und Nachnutzbarkeit von Forschungsdaten als wichtigen Baustein offener Wissenschaft ansieht. Empirisch scheint der dritte Motivationstyp bei den Promovierenden die derzeit dominante Form zu sein.

9. *Forschungsdatenpublikation oft unabhängig von Dissertationsschrift*

Entgegen der Ausgangsannahme, dass die Dissertationsschrift und die zugehörigen Forschungsdaten gleichzeitig publiziert werden, zeigt sich, dass besonders dann, wenn die Datenpublikation nicht von vornherein im Publikationsplan vorgesehen ist, der Abschluss der Promotion zunächst gegenüber einer Forschungsdatenpublikation priorisiert wird. Daher wird oft eine möglichst frühzeitige Veröffentlichung der Dissertationsschrift angestrebt. Aufbereitung und Publikation der Forschungsdaten werden eher für den Zeitraum nach der Titelzuerkennung angestrebt.

10. *Migration der Forschungsdaten von Fremdrepositorien ggf. erwünscht*

Es gibt weiterhin Fälle, bei denen Forschungsdaten bereits vor Abschluss der Arbeit auf externen Plattformen veröffentlicht wurden. Entweder sollten dadurch Nachnutzungsszenarien frühzeitig realisiert werden oder die Forschungsdaten sind zugleich in übergreifende Zusammenhänge wie kooperative Forschungsprojekte eingebunden. Hier besteht das Interesse, diese bereits extern veröffentlichten Daten möglichst in direkter Verbindung zur Dissertationsschrift nochmals zu veröffentlichen. Es wird auch betont, dass der natürliche Erscheinungsort einer Forschungsdatenpublikation zur Promotion die jeweilige Hochschule ist.

11. *Kontrolle über Weitergabe eigener Forschungsdaten statt Open Research Data*

Viele Promovierende möchten selbst entscheiden, wem und unter welchen Bedingungen sie ihre Forschungsdaten zur Einsicht und Nachnutzung überlassen. Ein Nachweis von Forschungsdaten durch Metadaten in Suchmaschinen bzw. Discovery Systemen gilt durchaus als wichtig, um die eigene Sichtbarkeit zu erhöhen. Die Freigabe der Daten selbst soll dagegen nur nach Rücksprache erfolgen.

12. *Dokumentation kaum standardisiert und sehr fachspezifisch*

Für die Dokumentation von Forschungsdaten existieren unterschiedliche disziplinäre Anforderungen, aber selten Standards bzw. Vorlagen. Hier existiert eindeutig ein Desiderat. Die Dokumentation von Forschungsdaten wird oft als erheblicher zusätzlicher Aufwand bewertet, der viele Promovierende von einer Forschungsdatenpublikation absehen lässt.

Zusammenfassung

Der erste Teil des Beitrages reflektierte mögliche Verantwortungsstrukturen für die Langzeitar Archivierung dissertationsbezogener Forschungsdaten. Aus dem Gesagten ergeben sich Eckpunkte für eine Forschungsdaten-Policy der DNB, die auch auf andere Publikationsformen erweiterbar ist. Hierbei sind insbesondere die unterschiedlichen Integrationsgrade von Forschungsdaten in

eine Dissertation (Supplement oder Referenz) und die Bedeutung des inhaltlichen Bezugs der Daten zur Dissertation als Selektionskriterium zu beachten.

Im zweiten Teil wurden Zwischenergebnisse der Anforderungsanalyse für die Archivierung und Veröffentlichung von dissertationsbegleitenden Forschungsdaten vorgestellt. Als vorläufiges Fazit lässt sich festhalten, dass bisher eher punktuelle Bedarfe bestehen. Die relative Randständigkeit des Themas in der Wissenschaftspraxis lässt vermuten, dass bestehende Initiativen eher wissenschafts- und förderpolitisch und damit extrinsisch zu den Fach-Communities motiviert sind. Insgesamt sollte also das zu entwickelnde Dienstleistungsspektrum neben den unabdingbaren technischen Rahmenbedingungen vor allem beratend ausgerichtet sein und die Möglichkeit enthalten, bei Bedarf auch eine direkte Begleitung auf Einzelfallebene anzubieten.

Literaturangaben

Arbeitskreis der Leiter wissenschaftlicher Rechenzentren in Baden-Württemberg (ALWR). 2017. „bwDataDiss - bwData für Dissertationen“. Online verfügbar unter <https://www.alwr-bw.de/kooperationen/bwdatadiss/>. Zuletzt geprüft am 13.08.2017.

Bundesministerium für Justiz und für Verbraucherschutz (BMJV). 2017. „Verordnung über die Pflichtablieferung von Medienwerken an die Deutsche Nationalbibliothek“. Online verfügbar unter <http://www.gesetze-im-internet.de/pflav/index.html>. Zuletzt geprüft am 13.08.2017.

Deutsche Forschungsgemeinschaft (DFG). 2013. *Vorschläge zur Sicherung guter wissenschaftlicher Praxis: Empfehlungen der Kommission "Selbstkontrolle in der Wissenschaft"; Denkschrift*. Ergänzte Auflage. Weinheim: Wiley-VHC. doi:10.1002/9783527679188.oth1.

Deutsche Nationalbibliothek (DNB). 2013. "ODE - Opportunities for Data Exchange". Last modified February 08, 2013. Online verfügbar unter <http://www.dnb.de/DE/Wir/Projekte/Archiv/ode.html>. Zuletzt geprüft am 13.08.2017.

Fenner, Martin et al. 2015. D2.1: *Artefact, Contributor, and Organisation Relationship Data Schema*. doi:10.5281/zenodo.30799.

Karlsruhe Institute of Technology (KIT). 2015. "Öffentlicher Abschlussbericht von bwFDM-Communities". Online verfügbar unter <http://bwfdm.scc.kit.edu/bwFDM-Communities.php>. Zuletzt geprüft am 13.08.2017.

Neuroth et al. 2009. *nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*. Version 2.0. Göttingen: Verlag Werner Hülsbusch,

Reilly, Susan, et al. 2011. *Opportunities of Data Exchange: Report on Integration of Data and Publications*. Online verfügbar unter [hdl:10013/epic.40198.d001](https://hdl.handle.net/10013/epic.40198.d001). Zuletzt geprüft am 13.08.2017.

RfII – Rat für Informationsinfrastrukturen. 2016. *Leistung aus Vielfalt: Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland*. Göttingen. Online verfügbar unter [urn:nbn:de:101:1-201606229098](https://nbn-resolving.org/urn:nbn:de:101:1-201606229098). Zuletzt geprüft am 13.08.2017.

Simukovic, Elena, Maxi Kindling, Peter Schirnbacher. 2013. „Umfrage zum Umgang mit digitalen Forschungsdaten an der Humboldt-Universität zu Berlin“. Online verfügbar unter <http://edoc.hu-berlin.de/oa/reports/reFIYMGduNiVE/PDF/22YavRASzVauc.pdf>. Zuletzt geprüft am 13.08.2017.

Tarazona Rua, Maria et al., eds., 2015. *Access to and Preservation of Scientific Information in Europe*. Luxemburg: Publications Office of the European Union. doi:10.2777/975917.

Tenopir et al. 2017. Research Data Services in European Academic Research Libraries. *LIBER Quarterly* 27(1), 23–44. <http://doi.org/10.18352/lq.10180>.