

o | bib

Das offene
Bibliotheksjournal



2

2018

V | D | B

Verein Deutscher
Bibliothekarinnen
und Bibliothekare

Inhaltsverzeichnis

Editorial: Forschungsdaten – Aufgabe und Herausforderung für Bibliotheken.....IV

Themenschwerpunkt

Leitlinie? Grundsätze? Policy? Richtlinie? Forschungsdaten-Policies an deutschen Universitäten	1
<i>Bea Hiemenz und Monika Kuberek, Universitätsbibliothek der Technischen Universität Berlin</i>	
HeFDI – Die landesweite Initiative zum Aufbau von Forschungsdateninfrastrukturen in Hessen... 14 <i>Ortrun Brand, Philipps-Universität Marburg</i> <i>Wolfgang Stille, Technische Universität Darmstadt</i> <i>Joachim Schachtner, Philipps-Universität Marburg</i>	
Von der Idee zum Konzept – Forschungsdatenmanagement an der Universität zu Köln	28
<i>Jens Dierkes, Universitäts- und Stadtbibliothek Köln</i> <i>Constanze Curdt, Regionales Rechenzentrum der Universität zu Köln</i>	
Datenmanagementpläne leicht gemacht: TUB-DMP, das Web-Tool der TU Berlin für Datenmanagementpläne.....	47
<i>Monika Kuberek und Fabian Fürste, Universitätsbibliothek der Technischen Universität Berlin</i>	
Offene Forschungsdaten an der Universität Heidelberg: von generischen institutionellen Repositorien zu fach- und projektspezifischen Diensten	61
<i>Jochen Apel, Universitätsbibliothek Heidelberg</i> <i>Fabian Gebhart, Universitätsrechenzentrum Heidelberg</i> <i>Leonhard Maylein, Universitätsbibliothek Heidelberg</i> <i>Martin Wlotzka, Universitätsrechenzentrum Heidelberg</i>	
Pflichtablieferung von Dissertationen mit Forschungsdaten an die DNB Anlagerungsformen und Datenmodell	72
<i>Dirk Weisbrod, Deutsche Nationalbibliothek (bis Jan. 2018, seitdem DDPF) Frankfurt am Main</i>	
Anwendung von OPUS für Forschungsdaten – das Beispiel eines außeruniversitären Forschungsinstituts.....	79
<i>Wolfgang Peters-Kottig, Kooperativer Bibliotheksverbund Berlin-Brandenburg</i>	
SARA-Dienst: Software langfristig verfügbar machen.....	92
<i>Franziska Rapp, Stefan Kombrink, Volodymyr Kushnarenko, Universität Ulm</i> <i>Matthias Fratz, Daniel Scharon, Universität Konstanz</i>	
Sozialwissenschaftliche Forschungsdaten langfristig sichern und zugänglich machen: Herausforderungen und Lösungsansätze	106
<i>Anja Perry, GESIS – Leibniz-Institut für Sozialwissenschaften, Köln</i> <i>Jonas Recker, GESIS – Leibniz-Institut für Sozialwissenschaften, Köln</i>	

- Feldnotizen und Videomitschnitte
Zum Forschungsdatenmanagement qualitativer Daten am Beispiel der ethnologischen Fächer.. 123
Matthias Harbeck, Universitätsbibliothek Humboldt-Universität zu Berlin
Sabine Imeri, Universitätsbibliothek Humboldt-Universität zu Berlin
Wjatscheslaw Sterzer, Universitätsbibliothek Humboldt-Universität zu Berlin
- Der Chimäre auf der Spur: Forschungsdaten in den Geisteswissenschaften142
Fabian Cremer, Max Weber Stiftung
Lisa Klaffki, DARIAH-DE / Herzog August Bibliothek Wolfenbüttel
Timo Steyer, Forschungsverbund Marbach Weimar Wolfenbüttel/Herzog August Bibliothek Wolfenbüttel

Aufsätze

- Open-Access-Monografien in deutschsprachigen Universitätsverlagen..... 163
Claudia Schober, Universitätsbibliothek Würzburg
- Zeitretrieval in der verbalen Inhaltserschließung181
Esther Scheven, Deutsche Nationalbibliothek, Frankfurt am Main
Michael Wening, Öffentliche Bibliothek, Remscheid

Tagungsberichte

- Bericht zum Workshop „Rechtliche Aspekte bei digitalen Forschungsdaten“ an der Europa-Universität Viadrina Frankfurt (Oder)193
Thomas Hartmann, Europa-Universität Viadrina Frankfurt (Oder)
Ben Kaden, Humboldt-Universität zu Berlin
Michael Kleineberg, Humboldt-Universität zu Berlin
- Konferenz und Erfahrungsaustausch
26 Fachreferentinnen und Fachreferenten geisteswissenschaftlicher Fächer treffen sich
in Bensberg (ZBIW-Seminar 26.-27.02.2018) 202
Dorothee Graf, Universitätsbibliothek Duisburg-Essen
- Neue Konzepte für die Informations- und Literaturversorgung?
Eine Zusammenfassung der Fortbildung „Fachinformationsdienste: Angebote und
Perspektiven der überregionalen Informationsversorgung“ 208
Rainer Plappert, UB Erlangen-Nürnberg

Berichte und Mitteilungen

- Research Information Systems – fit for the future?
A report on the situation and plans of the University of Sheffield Library214
Tracey Clarke, Andy Bussey, The University of Sheffield Library
- Bericht aus der 73. Sitzung der Arbeitsgemeinschaft der Verbundsysteme am 15. und 16.
November 2017 in München..... 222
Edith Röschlau, Deutsche Nationalbibliothek
- Neue Verbundangebote für die Wissenschaft
Bericht der Arbeitsgemeinschaft der Verbundsysteme über ihre Planungen..... 229

Aus dem VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare e.V.

Vorstand und Vereinsausschuss

Kathrin Drechsel als Vorsitzende des Regionalverbands Sachsen – Sachsen-Anhalt – Thüringen verabschiedet.....	234
Ausschreibung der Kommissionen des VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare e. V. zum 1.10.2018	235
Gemeinsame Kommissionen von VDB und dbv – Mitglieder der neuen Amtszeit	240
Bibliothek – Qualifikation – Perspektiven: D-A-CH-S-Tagung der bibliothekarischen Verbände 13. bis 14. Februar 2019, Ludwig-Maximilians-Universität München.....	241

Kommissionen

Kommission für Rechtsfragen Tätigkeitsbericht 2017	244
<i>Thomas Witzgall, Thüringer Universitäts- und Landesbibliothek Jena (Vorsitzender der Kommission)</i>	
Gemeinsame Baukommission des Deutschen Bibliotheksverbands und des Vereins Deutscher Bibliothekarinnen und Bibliothekare Jahresbericht 2017	245
<i>Susanne Kandler, Universitätsbibliothek der TU Freiberg (Vorsitzende der Kommission)</i>	

Landes- und Regionalverbände

Regionalverband Berlin – Brandenburg Jahresbericht 2017	247
<i>Janin Präßler, Stadtbibliotheken Treptow-Köpenick, Berlin (Vorsitzende des Regionalverbandes)</i>	
Landesverband Hessen Hessischer Bibliothekstag 2018, 14. Mai 2018 in Wetzlar	250
<i>Julijana Nadj-Guttandin, Deutsche Nationalbibliothek (Schriftführerin des Landesverbandes Hessen)</i>	
Regionalverband Sachsen – Sachsen-Anhalt – Thüringen Mitgliederversammlung: Wahl eines neuen Vorstandes.....	253
<i>Susann Özüyman, ULB Sachsen-Anhalt (Stellvertretende Vorsitzende des Landesverbandes Sachsen – Sachsen-Anhalt – Thüringen)</i>	
Landesverband Bayern Eine Zusammenfassung der Fortbildung „Fachinformationsdienste: Angebote und Perspektiven der überregionalen Informationsversorgung“	254

Personalia

Aus der Mitgliederverwaltung.....	255
-----------------------------------	-----

Editorial

Forschungsdaten – Aufgabe und Herausforderung für Bibliotheken

Liebe Leserinnen, liebe Leser,

spätestens seit den Empfehlungen des Rates für Informationsinfrastrukturen (RfII) im Jahr 2017 und der von Bund und Ländern formulierten Perspektive der Einrichtung einer Nationalen Forschungsinfrastruktur (NFDI) im Jahr 2018 steht das Thema Forschungsdaten ganz oben auf der bibliothekarischen Agenda und markiert zugleich einen sichtbaren Umbruch in wesentlichen bibliothekarischen Handlungsfeldern – auch wenn man dies auf den ersten Blick vielleicht gar nicht wahrnehmen mag. Daten, die im Forschungsprozess entstanden oder gesammelt worden sind, zu archivieren, zu verzeichnen und zugänglich zu machen, ist eine relative neue Aufgabe und zwingt Bibliotheken dazu, Arbeitsfelder zu verändern und Geschäftsgänge an die neuen Gegebenheiten anzupassen. Deshalb war es für das Herausbergremium von o-bib naheliegend, dieses wichtige und neuartige Feld in einem Themenschwerpunkt näher zu durchdringen und von unterschiedlichen Seiten zu beleuchten.

Schon die schiere Vielfalt der existierenden Forschungsdaten und ihre disziplinären Besonderheiten werfen fundamentale Fragen auf. Wenn nach der breiten Definition der Allianz der deutschen Wissenschaftsorganisationen Forschungsdaten Daten sind, die „im Zuge wissenschaftlicher Vorhaben z.B. durch Digitalisierung, Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen“, dann fallen darunter Vermessungsdaten, digitale Karten oder Fotografien, Volltexte, Datenbanken, 3-D-Rekonstruktionen, Filme und Videos, Vektorzeichnungen, Spektrogramme, aufbereitete und angereicherte Daten oder Dokumente, digitale Anmerkungen oder Bibliografien. Diese Aufzählung könnte beliebig fortgeführt werden. Entscheidend ist, dass es sich um digital verarbeitbare Daten handelt, die gleichsam den Rohstoff einer digital arbeitenden Wissenschaft bilden. Oft wird dabei übersehen, dass auch die Publikationen selbst und nicht nur die Publikationen begleitenden Daten zu Forschungsdaten werden, wenn sie entsprechend aggregiert und ausgelesen werden können. Auch daran wird sichtbar, wie weitreichend der derzeit stattfindende Paradigmenwechsel ist. Nach dem Ende der analogen Sammlung müssen sich Bibliotheken mit dem Aufbau einer neuen digitalen Sammlung befassen, um etwa Text-and-Data-Mining (TDM) zu ermöglichen.

Damit einher gehen Forderungen nach einem neuen Umgang mit den von der öffentlichen Hand bezahlten Forschungsergebnissen. Die vielzitierten FAIR-Prinzipien spielen dabei ebenso eine Rolle wie die sich mehr und mehr durchsetzende Einsicht, dass moderne Forschung ohne Open Access eigentlich undenkbar ist. Schließlich ergeben sich in der Organisation und Verwaltung der Forschungsdaten zahlreiche, oft schwierige Detailprobleme. Auch wenn die abstrakten Modelle des Research Data Life Cycle mittlerweile den meisten bekannt sein dürften, so ist doch die konkrete Umsetzung gemäß den Anforderungen der Wissenschaft, die Verwaltung der Daten, ihre Dokumentation, Langzeitarchivierung und persistente Zugänglichkeit immer noch eine große Herausforderung und deshalb weiter Gegenstand eigener bibliothekarischer und infrastruktureller Forschungsprojekte.

Im Themenschwerpunkt des vorliegenden Heftes wollen wir uns des Themas in seinen vielfältigen Facetten annehmen. Die Beiträge stammen aus verschiedenen Disziplinen, von unterschiedlichen Protagonisten und greifen vor allem die mit Forschungsdaten verbundenen praktischen Probleme auf – von publikationsbegleitenden Forschungsdaten über Forschungsdatenmanagementpläne und Repositorien bis hin zu Softwarearchivierung. Zusätzlich zu den Fachaufsätzen finden sich in diesem Heft auch in den Rubriken „Tagungsberichte“ und „Berichte und Mitteilungen“ einschlägige Texte aus dem Bereich Forschungsdaten.

Dass das Thema von zentraler Relevanz ist, zeigt sich auch in dem Umstand, dass wir nach dem Call for Papers für diesen Themenschwerpunkt förmlich von Beitragsvorschlägen überrollt wurden und deutlich mehr qualitätvolle Aufsätze eingereicht wurden, als in einem Heft verarbeitet werden konnten. Wir haben uns daher entschlossen, den Themenschwerpunkt auf zwei Hefte aufzuteilen. Freuen Sie sich also schon jetzt auch auf das nachfolgende Heft 3/2018 von o-bib, in dem Sie weitere Beiträge zum Thema Forschungsdatenmanagement finden werden.

Für das o-bib-Team

Thomas Stäcker und Helge Steenweg

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2SIV-V>

Leitlinie? Grundsätze? Policy? Richtlinie?

Forschungsdaten-Policies an deutschen Universitäten

Bea Hiemenz und Monika Kuberek, Universitätsbibliothek der Technischen Universität Berlin

Zusammenfassung:

Die Anzahl der Forschungsdaten-Policies an deutschen Universitäten ist seit 2014 deutlich gestiegen und die Kurve geht kontinuierlich nach oben. Eine Besonderheit der deutschen Policies ist auf den ersten Blick ersichtlich: Es gibt keine einheitliche Bezeichnung – verwendet werden „Leitlinie“, „Grundsätze“, „Policy“, „Richtlinie“. Auch zeigen die Policies deutliche Unterschiede auf, was Umfang und Inhalte angeht. Um die Forschungsdaten-Policies an deutschen Universitäten weiter zu befördern, entwickelt die Technische Universität Berlin seit August 2017 im BMBF-Projekt „Modalitäten und Entwicklung institutioneller Forschungsdaten-Policies“ entsprechende Handlungsanleitungen. Ziel der ersten Projektphase ist die Konzeptionierung und Erstellung eines „Baukastens“ für institutionelle Forschungsdaten-Policies („Policy-Kit“). Als methodischer Ansatz werden Forschungsdaten-Policies deutscher Hochschulen gesammelt und evaluiert und mit internationalen Empfehlungen zu Forschungsdaten-Policies abgeglichen. Die Ergebnisse werden in diesem Artikel vorgestellt.

Summary:

The number of research data policies at German universities has increased significantly since 2014 and there is a continuous trend upward. At first glance, a special characteristic of the German policies is evident: There is no single term, but there are several terms in use – “Leitlinie”, “Grundsätze”, “Policy”, “Richtlinie”. The policies also show significant differences regarding scope and content. In the course of the BMBF project “Modalities and development of institutional research data policies”, the Technische Universität Berlin is developing instructions to further promote the research data policies at German universities. The goal of the first phase of the project is the design and creation of a “construction kit” for institutional research data policies (“Policy Kit”). As a methodological approach, research data policies of German universities are collected and evaluated, and compared to international recommendations on research data policies. The results are presented in this article.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S1-13>

Autorenidentifikation: Kuberek, Monika: ORCID: <https://orcid.org/0000-0002-1672-5271>;

Hiemenz, Bea: ORCID: <https://orcid.org/0000-0002-7327-9415>

Schlagwörter: Forschungsdaten-Policy; FD-Policy; Forschungsdaten-Leitlinie; Forschungsdatenmanagement; FAIR Data Principles

1. Aktuelle Situation

Die Anzahl der Forschungsdaten-Policies (FD-Policies) an deutschen Universitäten ist seit 2014 deutlich gestiegen und die Kurve geht kontinuierlich nach oben. Ein wichtiger Impulsgeber war die Hochschulrektorenkonferenz (HRK) in den Jahren 2014 und 2015 mit ihren beiden Empfehlungen zum Forschungsdatenmanagement (FDM). In der Empfehlung 2014 forderte sie die Hochschulleitungen

auf, die Steuerung der strategischen Prozesse des FDMs in die Hand zu nehmen und „Leitlinien zum Umgang mit digitalen Forschungsdaten abzustimmen“.¹ 2015 bekräftigte sie diesen Ansatz: „Als bewährtes Mittel, das Thema FDM in der Hochschule ganz oben auf die Agenda zu setzen, hat sich die Veröffentlichung einer sogenannten ‚Forschungsdatenpolicy‘ im Sinne einer strategischen Leitlinie für das FDM erwiesen.“² Bis Ende 2017 haben 22, das ist rund ein Fünftel der insgesamt 106 Universitäten in Deutschland, eine FD-Policy veröffentlicht. Auch wenn diese Zahl im internationalen Vergleich zunächst gering erscheint, beispielsweise im Vergleich zu Großbritannien, wo aktuell drei Fünftel der Universitäten eine FD-Policy veröffentlicht haben,³ ist doch aufgrund der aktuellen Bereitschaft, sich mit dem Thema zu befassen und der jährlichen Steigerung der Zahlen damit zu rechnen, dass die relativen Zahlen in Deutschland bald dem internationalen Vergleich standhalten werden.

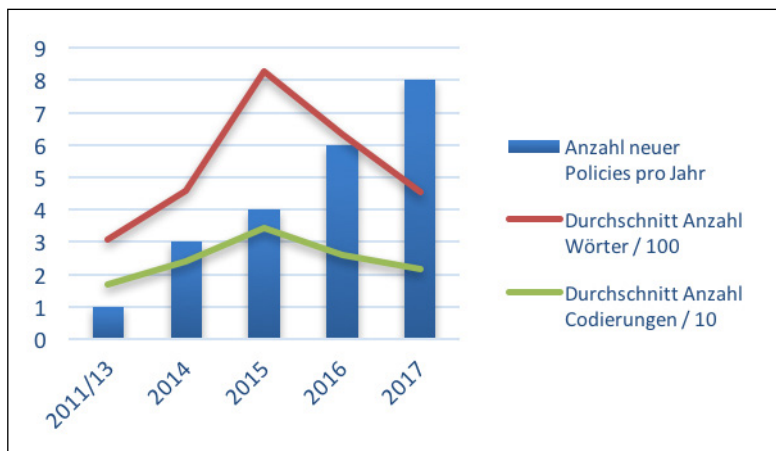


Abb. 1: Anzahl verabschiedeter Policies pro Jahr und Anzahl der Wörter und Codierungen der Policies

Ein Phänomen, das augenscheinlich ist: Schaut man sich die rein zahlenmäßige Auswertung der Wörter an, so ist abzulesen, dass die FD-Policies sehr stark im Umfang variieren und die umfangreichsten Policies 2015 verabschiedet wurden (s. Abb. 1). Ob dieses Phänomen mit den Empfehlungen der Hochschulrektorenkonferenz zusammenhängt, sei dahingestellt. Sicher ist, dass die Empfehlungen zum Anlass genommen wurden, sich an den Universitäten verstärkt mit dem FDM und der Erstellung von FD-Policies zu beschäftigen.

- 1 „Management von Forschungsdaten – eine zentrale strategische Herausforderung für Hochschulleitungen: Empfehlung der 16. Mitgliederversammlung der HRK am 13. Mai 2014 in Frankfurt am Main,“ HRK, 2, zuletzt geprüft am 31.01.2018, https://www.hrk.de/fileadmin/_migrated/content_uploads/HRK_Empfehlung_Forschungsdaten_13052014_01.pdf.
- 2 „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können. Orientierungspfade, Handlungsoptionen, Szenarien: Empfehlung der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel,“ HRK, 6, zuletzt geprüft am 22.01.2018, https://www.hrk.de/fileadmin/_migrated/content_uploads/Empfehlung_Forschungsdatenmanagement_final_Stand_11.11.2015.pdf.
- 3 Open Research Data Taskforce, *Research Data Infrastructures in the UK: Landscape Report* (30. Juni 2017), 17, zuletzt geprüft am 29.01.2018, <http://www.universitiesuk.ac.uk/policy-and-analysis/research-policy/open-science/Documents/ORDTF%20report%20nr%201%20final%2030%2006%202017.pdf>.

Um FD-Policies an deutschen Universitäten weiter zu befördern, entwickelt die Technische Universität Berlin (TU Berlin) im Teilprojekt „Modalitäten und Entwicklung institutioneller Forschungsdaten-Policies“ des BMBF-Verbundprojekts FDMentor⁴ entsprechende Handlungsanleitungen. Ziel bis zur Halbzeit des Projekts ist die Konzeptionierung und Erstellung eines „Baukastens“ für institutionelle FD-Policies („Policy-Kit“). Als methodischer Ansatz werden FD-Policies deutscher Universitäten gesammelt und evaluiert und mit internationalen Empfehlungen zu FD-Policies abgeglichen.

Eine strukturierte Inhaltsanalyse der FD-Policies gibt einen ersten Aufschluss darüber, welche inhaltlichen Kriterien für die Institutionen bei der Erarbeitung ihrer Policies von Bedeutung sind. In einem Abgleich mit internationalen Empfehlungen werden dann Unterschiede und Gemeinsamkeiten der inhaltlichen Priorisierung und konkreten Ausgestaltung der FD-Policies beleuchtet. In diesem Zusammenhang wurde ein Experteninterview mit Paolo Budroni geführt,⁵ dem Leiter des Arbeitspakets „Policy Development and Alignment“ (WP3) in dem H2020-Projekts LEARN,⁶ in dessen Rahmen auf internationaler Ebene ein Forschungsdatenmanagement (FDM)-Policy-Kit erstellt wurde, das von europäischen Institutionen nachgenutzt wird. Die systematische Untersuchung der deutschen FD-Policies und die Hinzuziehung internationaler Empfehlungen dienen als Basis zur Erstellung des Policy-Kit für deutsche Institutionen.

In dem vorliegenden Artikel werden erste Projektergebnisse vorgestellt.

2. Warum sollten Universitäten eine Forschungsdaten-Policy haben?

Als Argumente für eine FD-Policy haben sich im Laufe der Untersuchung insbesondere zwei Aspekte herauskristallisiert: Open Access von Forschungsdaten sowie Kosten- und Ressourcensteuerung.

Alle FD-Policies haben den Grundgedanken gemeinsam, dass öffentlich finanzierte Forschung auch als öffentliches Gut betrachtet und durch die getroffenen Aussagen und Regeln die beste Forschungspraxis angestrebt wird, von der sowohl die Forschungscommunity als auch die Gesellschaft profitieren sollen. Damit hat auch in den Universitäten der Gedanke des Open Access für Forschungsdaten Fuß gefasst – sie stellen eine wertvolle Ressource dar, die offen zugänglich und nachnutzbar sein sollte. In Deutschland hat sich die Allianz der deutschen Wissenschaftsorganisationen bereits 2010 für den „grundsätzlich offenen Zugang zu Daten aus öffentlich geförderter Forschung“⁷ ausgesprochen. Stark propagiert wurde der Gedanke 2010 auch auf europäischer Ebene, als die High Level Expert Group on Scientific Data der Europäischen Kommission ihre Vision zu Forschungsdaten im Jahre 2030 veröffentlichte: „Our vision is a scientific e-infrastructure that supports seamless access, use, re-use,

4 Das BMBF-Verbundprojekt „FDMentor/Kooperative Erarbeitung generalisierbarer Strategien und Lösungen für das Forschungsdatenmanagement unter Einbeziehung bestehender Expertise an universitären Zentraleinrichtungen“ wird seit 01.05.2017 gefördert und läuft bis 30.04.2019. Weitere Informationen auf der Projekt-Website <http://www.forschungsdaten.org/index.php/FDMentor>.

5 Bea Hiemenz, Institutionelle Forschungsdaten-Policies: Interview mit Paolo Budroni (Berlin: Technische Universität, 2018), <http://dx.doi.org/10.14279/depositononce-6804>.

6 Weitere Informationen zum Projekt LEARN, zuletzt geprüft am 13.02.2018, learn-rdm.eu.

7 „Grundsätze zum Umgang mit Forschungsdaten,“ Allianz der deutschen Wissenschaftsorganisationen, 24. Juni 2010, zuletzt geprüft am 14.02.2018, <https://www.allianzinitiative.de/de/archiv/forschungsdaten/grundsaeetze.html>.

and trust of data.”⁸ Auch das EU-Rahmenprogramm Horizon 2020 hat diesen Ansatz seit 2014 mit seinem Open Research Data Pilot⁹ intensiv befördert und damit auch an deutschen Universitäten wesentliche Anreize für die Veröffentlichung von Forschungsdaten gesetzt.

Neben dem Open Access für Forschungsdaten gibt es einen zweiten zentralen Aspekt, der für die Verabschiedung einer Policy spricht: Die Policy kann als zentrales Element bei der Kosten- und Ressourcensteuerung im Bereich FDM fungieren, wie Budroni es im Interview klar formuliert.¹⁰ Bei der Erstellung und Verabschiedung einer Policy muss sich die Institution zwangsläufig damit beschäftigen, wie sie die Aufgabe FDM konzipiert und angeht. Im Fokus steht hierbei die Frage nach den personellen, organisatorischen und technischen Kapazitäten, die für das FDM benötigt werden. Dieser zweite Aspekt findet in den deutschen Universitäten insofern seinen Niederschlag, als in der Hälfte der FD-Policies ein Datenmanagementplan, der auch Aussagen zu den Kosten des FDM enthalten sollte, als wichtig erachtet wird.

Alle weiteren Vorteile bzw. Gründe, eine Policy zu veröffentlichen, ergeben sich hieraus: sei es eine höhere Reputation der Universität, die Transparenz für Hochschulangehörige, die sich klare Regelungen zum FDM wünschen¹¹ oder Vorteile bei der Mitteleinwerbung durch die zunehmende Bedeutung der Sicherung, Aufbewahrung und nachhaltigen Verfügbarkeit von Forschungsdaten bei nationalen und internationalen Förderorganisationen.

3. Leitlinie, Grundsätze, Policy oder Richtlinie?

Eine Besonderheit der deutschen FD-Policies ist auf den ersten Blick ersichtlich: Es gibt keine einheitliche Titelgebung. Die Hälfte der Universitäten wählt die Bezeichnung „Leitlinie“, rund ein Viertel „Grundsätze“, gefolgt von „Policy“; die wenigsten benutzen den Begriff „Richtlinie“ (s. Abb. 2). Auch in den Empfehlungen der wissenschaftspolitischen Akteure in Deutschland werden mehrere Bezeichnungen verwendet, teilweise analog. So benutzt etwa die HRK in ihren Empfehlungen 2014 den Begriff „Leitlinie“,¹² in ihren Empfehlungen 2015 sowohl „Leitlinie“ als auch „Policy“.¹³ Die DFG hat 2015 „Leitlinien zum Umgang mit Forschungsdaten“¹⁴ herausgegeben und verweist auf ihrer

8 High Level Expert Group of Scientific Data, *Riding the Wave: How Europe can gain from the rising tide of scientific data*, October 2010, 4, zuletzt geprüft am 14.02.2018, http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=707.

9 Die Regelungen sind auf der Horizon 2020-Website „Open Access & Data management“ übersichtlich dargestellt, mit Hinweisen zu den einschlägigen Dokumenten und Förderanträgen, zuletzt geprüft am 14.02.2018, http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm.

10 Siehe Hiemenz, 3 und 7.

11 Bruno Bauer et al., *Forschende und ihre Daten: Ergebnisse einer österreichweiten Befragung – Report 2015*, Version 1.2“ (Zenodo, 13. Oktober 2015), 59, <https://doi.org/10.5281/zenodo.32043>. Nach dieser Befragung von Forschenden im Rahmen von e-infrastructures austria erwarten über die Hälfte der Befragten (53%) die Veröffentlichung von Leitlinien oder Policies zum Umgang mit FD.

12 Siehe „Management von Forschungsdaten,“ HRK, 2 und 4.

13 Siehe „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können,“ HRK, 5, 6 und 18 (Leitlinie), 6 und 15 (Policy).

14 Deutsche Forschungsgemeinschaft, *Leitlinien zum Umgang mit Forschungsdaten* (2015), zuletzt geprüft am 14.02.2018, www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf.

Webseite „Umgang mit Forschungsdaten“¹⁵ auf diese wie auch auf die „Grundsätze zum Umgang mit Forschungsdaten“ der Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen.¹⁶

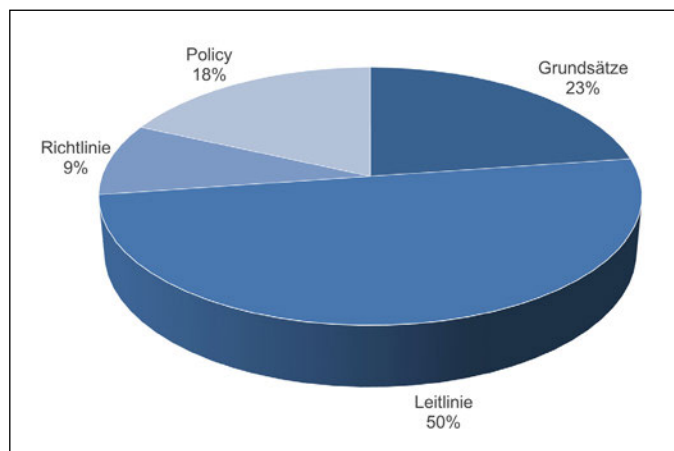


Abb. 2: Bezeichnungen der FD-Policies

Im Rahmen des Projekts der TU Berlin und in diesem Artikel verwenden wir den Begriff Policy. Die Suche nach einer Begriffsbestimmung von „Policy“ führt zunächst zu englischsprachigen Dictionaries. Das Oxford Dictionary stellt die Handlungsanleitung in den Vordergrund: „A course or principle of action adopted or proposed by an organization or individual“.¹⁷ Das Cambridge Dictionary fokussiert auf die gemeinsame Festlegung eines Plans, der vorgibt, was in bestimmten Situationen gültig ist: „A set of ideas or a plan of what to do in particular situations that has been agreed officially by a group of people, a business organization, a government, or a political party.“¹⁸ Gemeinsam ist beiden Formulierungen, dass übergeordnete Prinzipien aufgestellt werden, die in einer bestimmten Situation für eine bestimmte Zielgruppe gültig sind. Gilt dies auch für Grundsätze, Leit- und Richtlinien? Laut Duden Bedeutungswörterbuch¹⁹ ist „Leitlinie“ ein „bestimmender Grundsatz, leitender Gesichtspunkt, richtungweisender Anhaltspunkt (für das Handeln)“. Der Begriff „Grundsatz“ hat die zwei Bedeutungen 1. „feste Regel, die jemand zur Richtschnur seines Handelns macht“ und 2. „allgemeingültiges Prinzip, das einer Sache zugrunde liegt, nach dem sie ausgerichtet ist, das sie kennzeichnet; Grundprinzip.“ Der Begriff „Richtlinie“ wird dort so definiert: „von einer höheren Instanz ausgehende Anweisung für jemandes Verhalten in einem bestimmten Einzelfall, in einer Situation, bei einer Tätigkeit o. Ä.“ Am ehesten fallen somit „Richtlinie“ und „Policy“ zusammen.

15 „Umgang mit Forschungsdaten: DFG-Leitlinien zum Umgang mit Forschungsdaten,“ zuletzt geprüft am 14.02.2018, http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/index.html.

16 Siehe „Grundsätze zum Umgang mit Forschungsdaten,“ Allianz der deutschen Wissenschaftsorganisationen.

17 Siehe Oxford Online Dictionary (2018), <https://en.oxforddictionaries.com/definition/policy>, Suchwort „Policy“.

18 Siehe Cambridge Online Dictionary, <https://dictionary.cambridge.org/de/>, Suchwort „Policy“.

19 Siehe Duden Bedeutungswörterbuch, <https://www.duden.de/suchen/dudenonline/>, Suchwörter „Leitlinie“, „Grundsatz“, „Richtlinie“.

Eine direkte Übersetzung des Begriffs „policy“ führt zu „Richtlinie“ und „Strategie“ und (seltener) zu „Grundsatz“. Umgekehrt führt die direkte Übersetzung von Leitlinie zu „guideline“, von „Grundsätze“ zu „principles“ und (seltener) zu „policies“, von „Richtlinie“ zu „directive“, „guideline“ und „policy“.²⁰

Dieser Weg, Bedeutungszuweisung und Übersetzung, führt augenscheinlich zu keinem Ziel.²¹ Weiterführend ist hier ein Blick auf den Zusammenhang, in dem FD-Policies zum Tragen kommen – so beispielsweise in der europäischen Projektförderung. In den (englischsprachigen) Anträgen sollte eindeutig sein, worum es sich handelt; dies ist nur bei der Verwendung von „Policy“ gegeben. Budroni weist auf diese Problematik ausdrücklich hin: „Stellen Sie sich vor, Sie nehmen an einer Ausschreibung teil für ein großes Projekt von Horizon 2020. Und nehmen wir an, es geht um Governance in der European Open Science Cloud. Eine Voraussetzung ist, dass jede Universität eine Policy hat. Jetzt schreiben Sie diesen Antrag auf Englisch, wie übersetzen Sie da Leitlinien? Leitlinien als Policy zu übersetzen, wäre eine Irreführung, denn übersetzt wird es mit Guidelines. Eine Policy sollte immer auch als Policy betitelt sein, weil sich dadurch Regeln ergeben.“²²

4. Analyse der institutionellen FD-Policies in Deutschland

Die inhaltliche Auswertung erfolgt mit der Methode der strukturierten Inhaltsanalyse, mit Hilfe des Tools atlas.ti. Grundlage der Analyse bildet ein Kategoriensystem (s. Abb. 3), das auf Basis einer ersten Excel-Auswertung der FD-Policies der deutschen Universitäten und der im LEARN-Projekt in der Muster-Policy²³ und den Guidelines²⁴ verwendeten Inhalte entwickelt und im Laufe der Analyse angepasst wurde. Es besteht aus den übergeordneten Kategorien 1 bis 6, die jeweils in einzelne inhaltliche Kriterien untergliedert sind. Insgesamt wurden 39 Kriterien, zu denen in den Policies Aussagen gemacht wurden, in die Auswertung einbezogen.

Anhand des Kategorienschemas wurden die den Kategorien und Kriterien entsprechenden Textstellen der Policies in atlas.ti markiert (codiert). Für die Auswertung wurden diese markierten Textstellen (Quotations) nach unterschiedlichen Kriterien zusammengefasst und gruppiert, um anhand der Häufigkeiten der codierten Wörter Strukturen und Inhalte der Policies zu ermitteln. Die Auswertung erfolgte in zwei Schritten: zunächst wurden die Häufigkeiten analysiert, also das Vorhandensein der Kriterien abgeprüft und gezählt, anschließend eine vertiefende inhaltliche Auswertung vorgenommen, in der untersucht wurde, welche Art von Aussagen in den Policies zu den verschiedenen Kriterien gemacht werden.

20 Siehe Linguee, <https://www.linguee.de/deutsch-englisch>.

21 Es wäre interessant, die rechtlichen Implikationen der deutschen Begriffe zu untersuchen bzw. ob ihnen nur in einem bestimmten Zusammenhang, z.B. bei Verwendung in einer Satzung, ein rechtlicher Status zukommt. Die Klärung dieser juristischen Frage liegt nicht im Fokus des Projekts der TU Berlin und würde seinen Rahmen sprengen.

22 Siehe Hiemenz, 6.

23 LEARN, „Model Policy for Research Data Management (RDM) at Research Institutions/Institutes“, in *LEARN Toolkit of Best Practices for Research Data Management* (LEARN, 3. April 2017), 133–136. https://doi.org/10.14324/000_learn.00.

24 LEARN, „Guidance for Developing a Research Data Management (RDM) Policy. In *LEARN Toolkit of Best Practices for Research Data Management* (LEARN, 3. April 2017), 137–140. https://doi.org/10.14324/000_learn.00.

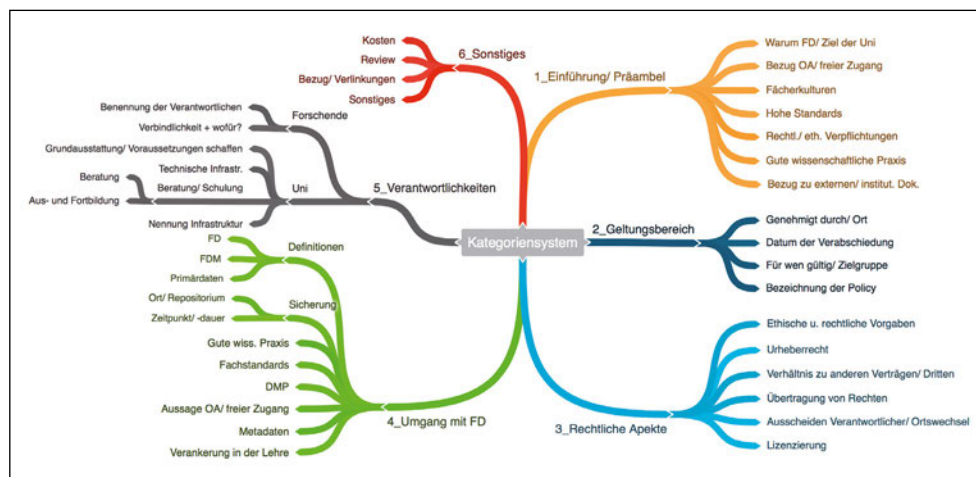


Abb. 3: Kategoriensystem zur Codierung der Forschungsdaten-Policies an deutschen Universitäten

Betrachtet man die Häufigkeit der inhaltlichen Kriterien, wird ein Grundkonsens aller Universitäten deutlich, was den Umgang mit Forschungsdaten und Verantwortlichkeiten angeht. So werden in allen Policies Forschende adressiert. Fast alle enthalten zudem Aussagen zur Beachtung ethisch-rechtlicher Aspekte beim FDM, der Fachstandards und zur Definition von Forschungsdaten sowie zur guten wissenschaftlichen Praxis und zum möglichst freien Zugang zu Forschungsdaten. In der Mehrzahl der Policies werden Beratung und Bereitstellung der technischen Infrastruktur bzw. die Ermöglichung des Zugangs hierzu als Verantwortlichkeit der Universität genannt. Weitere häufige Inhalte beziehen sich auf Zeitpunkt und Dauer der Speicherung, Zugang zu Forschungsdaten und die Erstellung eines Datenmanagementplans. Darüber hinaus nimmt eine Mehrzahl von 17 Universitäten Bezug auf weitere externe Dokumente. Dabei handelt es sich größtenteils um interdisziplinäre Empfehlungen, wie zum Beispiel die Empfehlungen der DFG; lediglich in vier Policies wird auf eigene ergänzende Handlungsempfehlungen zum FDM verwiesen. Und nur sechs Policies gibt es auch in einer englischen Version.

Die inhaltliche Auswertung ergibt dann ein differenzierteres Bild darüber, wie konkret die Universitäten ihre Policies in einzelnen Punkten ausformulieren und welche Wertigkeit sie den einzelnen Kriterien beimessen. Die gesamte Auswertung würde den Rahmen dieser Veröffentlichung sprengen, daher hier lediglich exemplarisch einige interessante Ergebnisse. Beispielsweise ist die „Beachtung der Fachstandards“ das meistgenannte Kriterium in der Kategorie *Umgang mit Forschungsdaten*. Allerdings ist die Diskrepanz zwischen Quantität und inhaltlicher Tiefe augenfällig. Generell üblich sind allgemeine Statements wie „Es sind fachspezifische Standards einzuhalten“; nur wenige Universitäten konkretisieren dies mit Aussagen wie „Die Fächer und Fakultäten können fachspezifische Richtlinien für typische datenintensive Forschungsvorhaben erstellen.“ Noch wenig Einzug in die FD-Policies gefunden hat der Aspekt, dass es geboten ist, nicht nur die Forschungsdaten, sondern auch die für die Auswertung verwendete Software zu speichern, ist sie doch für die Nachvollziehbarkeit und

Nachnutzbarkeit der Forschungsergebnisse wesentlich. Aktuell verweisen lediglich drei FD-Policies auf die Dokumentation der verwendeten Software, wobei eine Universität empfiehlt, neben den Forschungsdaten auch die zugehörige Software mit einer wissenschaftlichen Publikation öffentlich zugänglich zu machen. Auch bei der Kategorie *Rechtliche Aspekte* ist die Diskrepanz zwischen Quantität und inhaltlicher Tiefe festzustellen. So beinhalten fast alle Policies Aussagen zur Einhaltung rechtlicher und ethischer Aspekte, wobei allerdings in rund einem Drittel nur der knappe Hinweis steht, dass diese zu beachten sind. Immerhin führen zwei Drittel die rechtlichen und ethischen Aspekte weiter aus, wenn auch hier wiederum mit großen Unterschieden in der Ausführlichkeit; eine gängige Beschreibung lautet beispielsweise: „Die Wissenschaftlerinnen und Wissenschaftler beachten beim Forschungsdatenmanagement die Einhaltung ethischer, datenschutz- und urheberrechtlicher oder geheimhaltungswürdiger Belange.“ Im Unterschied zu diesen Beispielen kommt der Kategorie *Verantwortlichkeiten* nicht nur quantitativ eine wichtige Rolle in den FD-Policies zu, sondern sie ist in vielen Fällen auch inhaltlich ausgestaltet: Fast alle Policies enthalten zu dem Punkt „Verantwortlichkeiten der Forschenden“ mehr als einen Satz und adressieren sowohl die Projektleitung als auch die Forschenden. Die inhaltlichen Kriterien, die den Forschenden zugewiesen werden, sind die Einhaltung von Fachstandards, von ethisch-rechtlichen Verpflichtungen sowie der Prinzipien guter wissenschaftlicher Praxis und den Vorgaben Dritter, zudem Bestimmung von Speicherzeitpunkt und -dauer. Zusammenfassend lässt sich festhalten, dass der Bereich *Verantwortlichkeiten* die Forschenden als die Personen adressiert, die für das FDM und die Einhaltung der damit verbundenen Regelungen verantwortlich sind. Die Kategorie *Verantwortlichkeiten* umfasst zudem auch die Rolle der Universitäten, die in allen Policies thematisiert wird: Am häufigsten genannt wird die Beratung zum FDM, seltener auch Schulungsangebote sowie die Infrastruktur zum FDM, teilweise konkret in Bezug auf ein institutionelles Repositorium zur Datenspeicherung, teilweise durch die Zusage, eine Grundausstattung an Forschungsdateninfrastruktur zu implementieren oder die Voraussetzung zur Erfüllung der Grundsätze der Policy zu schaffen.

Die strukturierte Inhaltsanalyse liefert einen guten Überblick über die Inhalte der FD-Policies an den deutschen Universitäten. Doch wie vollständig sind sie? Gibt es relevante Dinge, die bislang in den FD-Policies nicht vorkommen oder zu kurz kommen? Zur Klärung solcher und ähnlicher Fragen werden die deutschen FD-Policies mit internationalen Empfehlungen zur Erstellung von FD-Policies abgeglichen.

5. Vergleich mit internationalen Empfehlungen

Für den Vergleich werden die folgenden Empfehlungen herangezogen: (1) die Muster-Policy,²⁵ die im LEARN-Projekt erarbeitet wurde, (2) die Anleitung zur Erstellung einer Forschungsdaten-Policy²⁶ aus dem LEARN-Projekt, (3) das Interview mit Budroni,²⁷ (4) die Five Steps der DCC²⁸, (5)

25 LEARN Toolkit 133–136.

26 Ebd. 137–140.

27 Siehe Hiemenz.

28 Martin Donnelly, *Five Steps to Developing a Research Data Policy: DCC 'Quickstart' Leaflets* (Edinburgh: Digital Curation Centre, 2014), zuletzt geprüft am 13.02.2018, <http://www.dcc.ac.uk/resources/policy-and-legal/five-steps-developing-research-data-policy/five-steps-developing-research>.

die Empfehlungen aus dem EU-Projekt RECODE²⁹ und (6) die Vorgaben von ANDS, dem Australian National Data Service³⁰. Der Vergleich wird anhand des bereits für die Analyse verwendeten Kriterienkatalogs (s. Abb. 3) durchgeführt. Im Folgenden sind beispielhaft einige relevante Ergänzungen für deutsche Policies aufgeführt, die sich aus den Empfehlungen eruieren lassen und bislang nicht in den Policies enthalten oder unterrepräsentiert sind.

LEARN empfiehlt, in der *Präambel* auch den Zusammenhang zum Forschungsverständnis der Institution herzustellen und eine Aussage darüber zu machen, nach welchem Standard die Institution das FDM betreiben will.³¹ Mögliche Ergänzungen in der Kategorie *Geltungsbereich/Gültigkeit* sind eine Definition des Anwendungsbereichs und eine Abgrenzung der Reichweite, die über die bisherigen Angaben in deutschen Policies hinausgeht: Für wen bzw. in welchen Fällen ist die Policy gültig und gibt es Ausnahmen? Dies impliziert auch die Frage, ob die Policy für alle Fachbereiche/Institute/Fakultäten und zum Beispiel auch für Studierende gültig ist, ob sie auch nicht-digitale Forschungsdaten abdeckt, ob sie jede Forschungsaktivität abdeckt oder beispielsweise nur extern geförderte.³²

In der Kategorie *Umgang mit Forschungsdaten* wird der freie Zugang zu Forschungsdaten in den untersuchten Empfehlungen stark propagiert, aber durchaus unterschiedlich gesehen: RECODE spricht sich dafür aus, dass Open Access verpflichtend ist, wobei die Forschungsdaten Nutzungsbedingungen beinhalten sollten.³³ Einen Mittelweg („SOLL“) zwischen Open Access als default („MUSS“) und einer reinen Empfehlung („KANN“) formuliert die Muster-Policy von LEARN: „In compliance with intellectual property rights, and if no third-party rights, legal requirements or property laws prohibit it, research data should be assigned a licence for open use.“³⁴ ANDS und LEARN empfehlen zudem, die Löschung von Forschungsdaten in den Policies zu thematisieren.³⁵ Dabei sollen sowohl ethische und rechtliche Aspekte als auch die Interessen von Dritten berücksichtigt werden und eine Dokumentation des Vorgangs erfolgen. Für Budroni ist die Löschung nur als Ausnahme denkbar: „Wenn die Daten einmal ins Repositorium eingestellt wurden, dürfen sie im Regelfall nicht mehr gelöscht werden. Sie können sie verstecken, Sie können sie verbergen, aber gelöscht werden dürfen die Daten nicht. Der zweite Aspekt bezieht sich auf das Recht des Vergessens: Da werden die Daten nicht gelöscht, sie werden inaktiv gesetzt. Das muss aber festgehalten werden. Wann darf ich etwas löschen? Ich darf löschen, wenn ich es muss: weil eine richterliche Anordnung kommt oder Inhalte gegen ein Gesetz verstoßen oder Inhalte der Institution schaden, pädophile Inhalte zum Beispiel – dann müssen sie gelöscht werden.“³⁶ Gegen das mögliche Löschen von Forschungsdaten sprechen zwei weitere Gründe: Im Normalfall werden die in einem Repositorium gespeicherten Forschungsdaten mit einem Persistent Identifier versehen, um sie im Internet eindeutig zu referenzieren und langfristig auffindbar zu machen. Dies

29 Victoria Tsoukala et al., *RECODE: Policy recommendations for open access to research data* (Zenodo, 27. April 2016), <https://doi.org/10.5281/zenodo.50863>.

30 „ANDS Guide, Institutional policies and procedures,“ ANDS, zuletzt geprüft am 13.02.2018, http://www.ands.org.au/_data/assets/pdf_file/0008/738782/Institutional-policies-and-procedures.pdf.

31 Siehe LEARN Toolkit, 138.

32 Siehe LEARN Toolkit, 140; Donnelly.

33 Siehe Tsoukala, 15.

34 Siehe LEARN Toolkit, 134.

35 Siehe ANDS Guide, 3; LEARN Toolkit, 134.

36 Siehe Hiemenz, 7.

bedingt entsprechende Verhaltensweisen der Institution: Sie kann die Nutzung einschränken oder den Zugriff sperren, aber die Forschungsdaten sollten nicht gelöscht werden. Eine Löschung von Forschungsdaten aus drittmittelfinanzierten Projekten kann zudem durch entsprechende Vorgaben der Fördergeber ausgeschlossen sein.

In Bezug auf die *rechtlichen Aspekte* stellen LEARN und ANDS die Rechteinhaberschaft in den Fokus, also die Klärung der Frage, wer welche Rechte an welchen Forschungsdaten hat. „Intellectual property rights (IPR) are defined in the work contract between a researcher and his or her employer. IPRs might also be defined through further agreements (e.g. grant or consortial agreements). In cases where the IPR belong to the institution that employs the researcher, the institution has the right to choose how to publish and share the data.“³⁷ Ein Verweis auf den Arbeitsvertrag oder auf sonstige Vereinbarungen fehlt bisher in den deutschen FD-Policies, einige FD-Policies enthalten Hinweise zum Urheberrecht.

In allen Empfehlungen ist hinsichtlich der *Verantwortlichkeiten* klar dargelegt, dass die Rollen und Verantwortlichkeiten sowie die Rechte und Pflichten sowohl für die Forschenden als auch für die Institution deutlich zu definieren sind. DCC empfiehlt, Institutionen, Projektleitungen, Wissenschaftlerinnen und Wissenschaftler sowie Mitarbeitende zu unterscheiden; nach ANDS sind auch „student researchers“ zu berücksichtigen.³⁸ Die Verantwortlichkeiten sollen gemäß LEARN nach Möglichkeit differenziert werden nach Forschenden, Förderern, Institution und forschungsunterstützenden Dienstleistern, denen jeweils ein spezieller Verantwortungsbereich zuzuweisen ist, so beispielsweise den Forschenden die Festlegung der Zugriffsrechte und der Lizenzen sowie Regelungen bei einem möglichen Ortswechsel. In der Verantwortlichkeit der Institutionen sieht LEARN neben Beratungs- und Schulungsangeboten insbesondere auch die Rechtsberatung für die Forschenden.³⁹

In der Kategorie *Sonstiges* sollte neben den Kosten, der Verlinkung zu weiteren institutionellen und externen Dokumenten insbesondere ein Review Eingang in die FD-Policies finden. Sowohl Budroni wie auch DCC befürworten ein Review, das regelmäßig oder nach Bedarf durchgeführt wird. RECODE empfiehlt darüber hinaus die Etablierung eines „Policy monitoring mechanism“, mit dessen Hilfe die Auswirkungen der Policy gemessen werden sollen, so dass in einem Review-Prozess Anpassungen möglich sind.⁴⁰ In Bezug auf die Kosten empfiehlt LEARN zu definieren, wer die Kosten des FDMs trägt, insbesondere auch Kosten, die nach Ende des Projekts anfallen. Budroni bemerkt darüber hinaus, dass die Kosten auch durch die Gestaltung der Policy beeinflusst werden können: „Die Kosten sind in der Policy impliziert. Zum Beispiel, wenn ich festlege: Nur Forschungsdaten, die im Rahmen eines EU-Forschungsprojektes entstehen, sind von der Policy betroffen. Das bedeutet dann, dass ich alle anderen Forschungsdaten ausschließe und somit den Kreis einenge und damit die Kosten senke.“⁴¹ Zur Verlinkung mit weiteren Dokumenten merkt Budroni an: „Eine Policy sollte

37 Siehe LEARN Toolkit, 133; ANDS Guide, 2.

38 Siehe Donelly; ANDS Guide, 2.

39 Siehe LEARN Toolkit, 6, 134f.

40 Siehe Hiemenz, 4; Donelly; Tsoukala, 33.

41 Siehe LEARN Toolkit, 140; Hiemenz, 7.

... in ihren Bestandteilen nicht zu viele Verweise auf weitere Dokumente haben, die einer ständigen Überprüfung unterliegen. Die Policy muss alleine für sich stehen können.⁴⁴²

Insgesamt zeigt der Vergleich mit den internationalen Empfehlungen, dass dort eine Fülle möglicher Ergänzungen für deutsche FD-Policies enthalten ist – sowohl hinsichtlich weiterer inhaltlicher Kriterien, wie auch hinsichtlich des Umfangs ihrer Ausgestaltung. Besonders augenfällig ist dies in Bezug auf die Aussagen zu einer offenen Zugänglichkeit der Forschungsdaten, die in internationalen Empfehlungen tendenziell verpflichtender gestaltet wird bis zur Default-Setzung des Open Access. Unterschiede zeigen die Empfehlungen auch bei der Frage der Einordnung der FD-Policy in das Gefüge der Policy-Landschaft an den Universitäten: Soll die FD-Policy alle Inhalte abdecken oder gibt es ergänzende Handlungsempfehlungen oder soll den Instituten oder Fakultäten empfohlen werden, die FD-Policy für ihr Fach zu spezifizieren? Während die internationalen Empfehlungen eher zu einer umfassenden institutionellen FD-Policy tendieren, enthalten einige deutsche Policies den Hinweis, dass die institutionelle FD-Policy durch die Fakultäten und/oder Institute fachspezifisch ausgestaltet werden kann/sollte.

6. Fazit und Ausblick

Als Ergebnis der Untersuchung kann festgehalten werden, dass in den FD-Policies der deutschen Universitäten ein Großteil der inhaltlichen Kriterien abgedeckt ist, die auch in den internationalen Empfehlungen thematisiert werden. Allerdings sind einige wesentliche Kriterien unterrepräsentiert, wie Kosten und die Löschung von Daten, Review-Verfahren für die Policies, die Klärung der Eigentümerrechte an den Forschungsdaten, Rechtsberatung und einige andere mehr. Hier können die Universitäten von der Erfahrung aus den EU-Projekten und den Ländern mit jahrelanger FDM-Praxis profitieren. Das von der TU Berlin erstellte Policy-Kit wird in einer Art Baukasten die in den deutschen FD-Policies identifizierten Kriterien sowie zusätzlich die in den internationalen Empfehlungen vorhandenen Kriterien als Textbausteine enthalten. Auf diese Weise sollen sich die Universitäten wie in einem modularen Baukasten bedienen und in ihrer FD-Policy die für sie relevanten Bausteine zusammenfügen können. Geplant ist, das Policy-Kit im ersten Halbjahr 2018 in einer ersten Version der Fachöffentlichkeit vorzustellen und zu diskutieren, bevor es dann im 3. Quartal 2018 veröffentlicht wird.

In der zweiten Halbzeit des Projekts wird die TU Berlin einen Policy-Leitfaden erarbeiten, der Strategien für FD-Policies von der Erstellung über die Diskussion mit den Stakeholdern bis zur Verabschiedung durch das Präsidium aufzeigt. Dazu werden Interviews mit Expertinnen und Experten herangezogen, d.h. Personen, die in ihrer Einrichtung für das FDM zuständig sind bzw. wesentlich am Erstellen der FD-Policy ihrer Einrichtung beteiligt waren.

42 Siehe Hiemenz, 4.

Literaturverzeichnis

- Allianz der deutschen Wissenschaftsorganisationen. „Grundsätze zum Umgang mit Forschungsdaten“, 24. Juni 2010. Zuletzt geprüft am 14.02.2018. <https://www.allianzinitiative.de/de/archiv/forschungsdaten/grundsaeetze.html>.
- Australian National Data Service. „ANDS Guide: Institutional policies and procedures“, 3. Februar 2017. Zuletzt geprüft am 13.02.2018. http://www.ands.org.au/__data/assets/pdf_file/0008/738782/Institutional-policies-and-procedures.pdf.
- Bauer, Bruno, Andreas Ferus, Juan Gorraiz, Veronika Gründhammer, Christian Gumpenberger, Nikolaus Maly, Johannes Michael Mühlegger, u. a. *Forschende und ihre Daten. Ergebnisse einer österreichweiten Befragung – Report 2015*. Version 1.2. Zenodo, Oktober 2015. <https://doi:10.5281/zenodo.32043>.
- DFG. *Leitlinien zum Umgang mit Forschungsdaten* (2015). Zuletzt geprüft am 14.02.2018. www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf.
- DFG. „DFG - Deutsche Forschungsgemeinschaft - Umgang mit Forschungsdaten“. Zuletzt geprüft am 14.02.2018. http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/antragstellung/nachnutzung_forschungsdaten/.
- Donnelly, Martin. *Five Steps to Developing a Research Data Policy: DCC ‘Quick-start’ Leaflets*. Edinburgh: Digital Curation Centre, 2014. Zuletzt geprüft am 13.02.2018. <http://www.dcc.ac.uk/sites/default/files/documents/publications/DCC-FiveStepsToDevelopingAnRDMpolicy.pdf>.
- European Commission. „Open access & Data management. Participant Portal H2020 Online Manual“. Zuletzt geprüft am 14.02.2018. http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm.
- Hiemenz, Bea. „Institutionelle Forschungsdaten-Policies: Interview mit Paolo Budroni“. Berlin: Technische Universität, April 2018. <http://dx.doi.org/10.14279/depositonce-6804>.
- High Level Expert Group on Scientific Data. *Riding the Wave: How Europe can gain from the rising tide of scientific data*. Oktober 2010. Zuletzt geprüft am 14.02.2018. http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=707.

- HRK. „Management von Forschungsdaten – eine zentrale strategische Herausforderung für Hochschulleitungen: Empfehlung der 16. Mitgliederversammlung der HRK am 13. Mai 2014 in Frankfurt am Main.“ Zuletzt geprüft am 31.01.2018. https://www.hrk.de/fileadmin/_migrated/content_uploads/HRK_Empfehlung_Forschungsdaten_13052014_01.pdf.
- HRK. „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können. Orientierungspfade, Handlungsoptionen, Szenarien: Empfehlung der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel.“ Zuletzt geprüft am 22.01.2018. https://www.hrk.de/fileadmin/_migrated/content_uploads/Empfehlung_Forschungsdatenmanagement__final_Stand_11.11.2015.pdf.
- LEARN. „Guidance for Developing a Research Data Management (RDM) Policy“. In *LEARN Toolkit of Best Practice for Research Data Management*, 137–140. LEARN, 2017. <https://doi.org/10.14324/000.learn.00>.
- LEARN. „LEARN Toolkit of Best Practice for Research Data Management“. LEARN, 3. April 2017. <https://doi:10.14324/000.learn.00>.
- LEARN. „Model Policy for Research Data Management (RDM) at Research Institutions/ Institut“. In *LEARN Toolkit of Best Practice for Research Data Management*, 133–136. LEARN, 2017. <https://doi.org/10.14324/000.learn.00>.
- Open Research Data Taskforce. *Research Data Infrastructures in the UK: Landscape Report*. 30. Juni 2017. Zuletzt geprüft am 29.01.2018. <http://www.universitiesuk.ac.uk/policy-and-analysis/research-policy/open-science/Documents/ORDTF%20report%20nr%201%20final%2030%2006%202017.pdf>.
- „*Riding the Wave: How Europe can gain from the rising tide of scientific data*. – Final report of the High Level Expert Group on Scientific Data“. A submission to the European Commission, Oktober 2010. Zuletzt geprüft am 14.02.2018. http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=707.
- Tsoukala, Victoria, Marina Angelaki, Vasso Kalaitzi, Bridgette Wessels, Lada Price, Mark J. Taylor, Rod Smallwood, u. a. *RECODE: Policy recommendations for open access to research data*. Zenodo, 27. April 2016. <https://doi:10.5281/zenodo.50863>.

HeFDI – Die landesweite Initiative zum Aufbau von Forschungsdateninfrastrukturen in Hessen

Ortrun Brand, Philipps-Universität Marburg

Wolfgang Stille, Technische Universität Darmstadt

Joachim Schachtner, Philipps-Universität Marburg

Zusammenfassung:

Digitale Forschungsdaten als Ressource sichern, aufbewahren und anderen zur Verfügung stellen: alle Phasen des Forschungsdatenzyklus adressiert das Projekt „Hessische Forschungsdateninfrastrukturen“ (HeFDI), in dem elf hessische Hochschulen und die HeBIS-Verbundzentrale unter Federführung der Philipps-Universität Marburg zusammenarbeiten. Das Hessische Ministerium für Wissenschaft und Kunst fördert dieses Vorhaben mit 3,4 Millionen Euro von 2016 bis 2020. Seit Oktober 2016 arbeitet ein Team von mehr als 20 Personen an den hessischen Hochschulen am Forschungsdatenmanagement, das heißt beispielsweise an Policies, Repositorien, Schulungen und unterstützenden Tools für das aktive Datenmanagement. Das Team berät zudem die Forschenden bedarfsorientiert zu ihren Datenmanagementstrategien entlang des gesamten Datenlebenszyklus und insbesondere bei der Einwerbung und Durchführung von Drittmittelprojekten. Ziel dieser Arbeit ist der Aufbau einer Forschungsdateninfrastruktur. Dazu werden die notwendigen organisatorischen Prozesse und technischen Entwicklungen zur Verankerung des Forschungsdatenmanagements – abhängig von den Anforderungen der beteiligten Hochschulen – angestoßen, koordiniert und etabliert. Das Projekt HeFDI nimmt dabei die strategisch bedeutsame Aufgabe der Sicherung und Bereitstellung von Forschungsdaten in einer koordinierten Gesamtstrategie an. Es vertritt einen Middle-out-Ansatz, bei dem Entscheidungs- und Arbeitsebenen systematisch miteinander verzahnt werden. Damit besitzt HeFDI Modellcharakter für institutionelle Kooperation und nachhaltige Governance-Strukturen im Bereich des Forschungsdatenmanagements; die in HeFDI erarbeiteten Lösungen stehen über Hessen hinaus exemplarisch für Synergie, Effizienz, Abstimmung und Arbeitsteilung.

Summary:

In a joint project called Hessian Research Data Infrastructures (HeFDI), eleven universities in the State of Hesse face the challenge to save digital research data as a resource, as well as to store and prepare digital research data for publication and re-use. HeFDI is funded by the Hessian Ministry for Science and the Arts with 3.4 million Euros, running from 2016 to 2020. Since October 2016, more than 20 experts are working together as a team to build up an infrastructure for research data management, e. g. by implementing policies, developing repositories, offering workshops and trainings and helping with tools for active data management. They also lend advice and support to researchers concerning data management strategies, especially with respect to third party funding at the early stage of project proposals. Thus, the HeFDI-project's objective consists in establishing a research data infrastructure which encompasses necessary organizational processes as well as technical developments and implementations. While there is strong cooperation, those processes and implementations are always adapted to local and institutional needs and the requirements of each university. Overall, HeFDI aims at tackling the crucial task to save and publish digital data by developing a coordinated common strategy for all project partners. At the same time, HeFDI represents

a middle out-strategy: decision making and action levels are closely intertwined, and cooperation is based on a reliable and functional governance system. Therefore, HeFDI serves as a role model for institutional cooperation and sustainable governance structures in research data management. Solutions developed within HeFDI reach out far beyond the federal state of Hesse as they can be taken as an example for possible synergies, efficiency, proper reconciliation and the distribution of tasks.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S14-27>

Autorenidentifikation: Brand, Ortrun: GND 104268233X, ORCID: <https://orcid.org/0000-0002-6850-5123>; Stille, Wolfgang: GND 136475523, ORCID: <http://orcid.org/0000-0003-4468-4208>; Schachtner, Joachim: GND 1160191484, ORCID: <http://orcid.org/0000-0001-7268-0699>

Schlagwörter: Forschungsdaten; FDM; Forschungsdatenmanagement

1. Einleitung

Das Management von Forschungsdaten gewinnt ohne Zweifel (wissenschafts-)politisch und infrastrukturell mehr und mehr an Bedeutung und gilt heute als eines der Zukunftsthemen, das die Wissenschaft flächendeckend umtreibt – von den Forschenden über Wissenschaftsinstitutionen und -organisationen bis hin zur Politik. Unabhängig von der Position im Wissenschaftssystem oder der jeweiligen fachlichen Perspektive sind sich im Grunde alle Akteure einig, dass vordringlich nachhaltige Strukturen geschaffen werden müssen, um das Management von Forschungsdaten in ihrem gesamten Lebenszyklus zu unterstützen und – in zunehmend auf digitalen Daten basierenden Forschungskulturen – die Zukunft des Wissenschaftsstandortes Deutschland dauerhaft zu sichern. Konsens besteht auch dahingehend, dass das Problem nur in einer *gemeinsamen* Anstrengung angegangen werden kann. Dies ist vor dem Hintergrund einer diversifizierten deutschen Forschungslandschaft mit autonomen Hochschulen, außeruniversitären Forschungseinrichtungen, unterschiedlichen Hochschulgesetzen und Finanzierungsmodellen eine nicht zu unterschätzende Herausforderung. Forschungsdatenmanagement – für sich selbst genommen Herausforderung genug – stellt freilich nur eine Facette in der allumfassenden Digitalisierung der Wissenschaft dar. Digitale Daten liegen im Zuge dieser Veränderungen heute in schnell wachsenden Quantitäten vor, sind durch das Internet weitgehend ortsunabhängig und potentiell jederzeit zugreifbar. Durch stetig wachsende Rechnerleistung und algorithmische Verfahren z.B. aus dem Bereich des maschinellen Lernens oder der mathematischen Simulation entstehen gänzlich neue Analyseformen und Methoden; sie versprechen, den wissenschaftlichen Erkenntnisgewinn auf rasante Weise voranzutreiben. Gleichzeitig wächst die Heterogenität der Daten. Letzteres eröffnet einerseits eine Vielzahl von Möglichkeiten für neue interdisziplinäre Forschungsfragen, die sich nun durch die Kombination von Daten aus unterschiedlichen Bereichen und geeigneten Verfahren bearbeiten lassen, und schafft andererseits Herausforderungen organisatorischer und technischer Art.

Dieser allgemeine gesellschaftliche und wissenschaftliche Umbruch erreicht die Hochschulen und ihre Infrastruktureinrichtungen. Bibliotheken und Rechenzentren befassen sich folglich zunehmend mit den Anforderungen des Forschungsdatenmanagements; sie bauen neue Angebote auf oder entwickeln bestehende Dienste weiter. Damit stellt sich die Frage, wie die Hochschulen den Prozess

unterstützen können und welche institutionelle Struktur, welche Governance und welche Modelle der Zusammenarbeit sich dabei bewähren. Gleichzeitig stellt sich die Frage, ob die derzeitig noch recht vielfältigen Entwicklungen in diesem Bereich koordiniert und arbeitsteilig vorangetrieben werden können und wie entsprechende Strukturen nachhaltig zu finanzieren sind.

Wir adressieren diese Fragen im vorliegenden Beitrag, indem wir die bundeslandweite Initiative *Hessische Forschungsdateninfrastrukturen (HeFDI)* vorstellen. Das Projekt HeFDI ist eines von bisher sechs geförderten großen kooperativen Infrastrukturprojekten in Hessen. Es wird gemeinsam von elf hessischen Hochschulen getragen und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) finanziert. Es ist Teil des 2012 gestarteten Prozesses der hessischen Landeshochschulentwicklungsplanung. Wir sehen in HeFDI ein Good Practice-Beispiel für gemeinsamen Aufbau, institutionelle Kooperation und nachhaltige Governance von Forschungsdateninfrastrukturen, das mit seinem systematischen Middle-out-Ansatz als Modell nicht auf Hessen beschränkt bleibt und in der aktuellen Diskussion weit über Hessen hinaus einen Beitrag leistet.

Dazu beschreiben wir im Folgenden zunächst die Entstehung und die Arbeitsbereiche von HeFDI (Abschnitt 2). Anschließend gehen wir auf die im Projekt etablierte Governance-Struktur sowie die institutionelle Arbeitsteilung ein (Abschnitt 3). In einem dritten Schritt beschreiben wir die Vernetzung auch über den Verbund hinaus sowie die daraus bereits erwachsenen Spill-over-Effekte und weitere Schritte (Abschnitt 4).

2. Entstehung und Arbeitsbereiche

Die Entstehung von HeFDI ordnet sich in wissenschaftspolitische und infrastrukturelle Entwicklungen ein und kommt somit „nicht von ungefähr“: Akteure wie die Allianz der Deutschen Wissenschaftsorganisationen, die Deutsche Forschungsgemeinschaft, die Hochschulrektorenkonferenz, die League of European Research Universities oder der Wissenschaftsrat haben in den vergangenen Jahren Empfehlungen verabschiedet, die das Forschungsdatenmanagement als ein zentrales Zukunftselement für die Entwicklung der Wissenschaft identifizieren.¹

Die beteiligten hessischen Universitäten und Hochschulen für angewandte Wissenschaften haben im Austausch mit dem HMWK diesen Prozess bereits 2012 durch eine hessenweite Arbeitsgruppe im

1 Allianz der Deutschen Wissenschaftsorganisationen, „Grundsätze zum Umgang mit Forschungsdaten,“ zuletzt geprüft am 06.12.2016, http://www.wissenschaftsrat.de/download/archiv/Allianz_Grundsaeetze_Forschungsdaten.pdf; DFG – Deutsche Forschungsgemeinschaft, „Leitlinien zum Umgang mit Forschungsdaten“, zuletzt geprüft am 07.12.2016, http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf; HRK – Hochschulrektorenkonferenz, „Hochschule im digitalen Zeitalter: Informationskompetenz neu begreifen – Prozesse anders steuern; Entschließung der 13. Mitgliederversammlung der HRK am 20. November 2012 in Göttingen“, zuletzt geprüft am 09.07.2018, https://www.hrk.de/fileadmin/_migrated/content_uploads/Entschliessung_Informationskompetenz_20112012_01.pdf; HRK – Hochschulrektorenkonferenz, „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können: Orientierungspfade, Handlungsoptionen, Szenarien; Empfehlung der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel“, , zuletzt geprüft am 09.07.2018, https://www.hrk.de/fileadmin/_migrated/content_uploads/Empfehlung_Forschungsdatenmanagement_final_Stand_11.11.2015.pdf; LERU League of European Research Universities, „Roadmap for Research Data“, zuletzt geprüft am 18.01.2017, <https://www.leru.org/publications/leru-roadmap-forresearch-data>; Wissenschaftsrat, „Empfehlungen zu wissenschaftlicher Integrität. Positionspapier“, zuletzt geprüft am 23.01.2017, <http://www.wissenschaftsrat.de/download/archiv/4609-15.pdf>.

Rahmen der Landeshochschulentwicklungsplanung angestoßen. Im Sommer 2013 hat diese Arbeitsgruppe ein Positionspapier zur „Zukunft der wissenschaftlichen Infrastruktur in Hessen bis 2020“ verabschiedet, welches Handlungsfelder in den Bereichen technologische Infrastruktur, Medienversorgung und Infrastruktur für organisatorische Prozesse ausweist, an denen sich die Hochschulentwicklung in Hessen bis 2020 orientieren soll: Von *Hochleistungsrechnen* über *Campus-Management* bis zur *Langzeitarchivierung* wurden zentrale Themen identifiziert, und darüber fand schließlich der Aufbau von Strukturen zum Umgang mit Forschungsdaten Eingang in den Hessischen Hochschulpakt 2016-2020. Es heißt dazu in dem Pakt:

„Die Hochschulen schaffen im Rahmen der Landeshochschulentwicklungsplanung eine übergreifende, den Anforderungen moderner Forschung und Lehre genügende *Informationsinfrastruktur*. Aus dem Innovations- und Strukturentwicklungsbudget wird das Land Mittel für Maßnahmen für gemeinsam genutzte Infrastruktur zur Verfügung stellen.“ (Land Hessen)²

Die Entscheidung des Landes, die hessischen Hochschulen im Rahmen des Hochschulpakts mit Mitteln aus dem Innovations- und Strukturentwicklungsbudget auszustatten, ist somit eine Bestätigung dafür, dass hier rechtzeitig und gemeinsam ein Thema aufgegriffen wurde, das die Zukunft aller Hochschulen wesentlich mitbestimmt.

Der Einstieg in den hessischen Hochschulpakt und die Bewilligung des Projekts durch das HMWK fielen zeitlich zusammen mit der wegweisenden Publikation „Leistung aus Vielfalt“ des Rats für Informationsinfrastrukturen (RfII),³ in der dieser unter anderem die Einrichtung einer Nationalen Forschungsdateninfrastruktur (NFDI) vorschlägt. Auch macht der RfII mit seinen jüngeren Diskussionspapieren⁴ deutlich, dass wir alsbald in einem ersten Schritt auf die Einrichtung einer fachwissenschaftlich orientierten NFDI hoffen können. Daraus wird dreierlei deutlich: Erstens nimmt die Politik den durch das Ausschöpfen dieses Potentials zu erwartenden Mehrwert für die Wissenschaft und die Gesellschaft mittlerweile sehr deutlich wahr. Zweitens ist zu erwarten, dass zeitnah und fortlaufend in den nächsten Jahren die finanzielle Förderung dieser Entwicklung angepasst wird. Drittens zeigt sich damit, dass HeFDI zu einem Zeitpunkt gestartet ist, der wesentlich durch Umbrüche im Bereich des Forschungsdatenmanagements geprägt ist.

Gerade aber die zu erwartenden Veränderungen durch eine NFDI zeigen, dass es umso wichtiger ist, übergreifende Strukturen mit lokalen zu ergänzen und beide miteinander zu verzahnen. Denn wie wir meinen, ist eine NFDI grundlegend zu untermauern durch Aufbau und Unterstützung dort, wo die Forschungsdaten zunächst initial entstehen: an den Hochschulen. Diese Meinung vertritt im Übrigen auch die Hochschulrektorenkonferenz: Forschungsdatenmanagement und -infrastrukturen

2 Land Hessen, „Hochschulpakt 2016-2020“ (2015), Wiesbaden.

3 RfII – Rat für Informationsinfrastrukturen, „Leistung aus Vielfalt: Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsmanagements in Deutschland“, zuletzt geprüft am 06.12.2016, <http://www.rfii.de/?wpdmdl=1998>.

4 RfII – Rat für Informationsinfrastrukturen, „Schritt für Schritt – oder: Was bringt wer mit? Ein Diskussionsimpuls zu Zielstellung und Voraussetzungen für den Einstieg in die Nationale Forschungsdateninfrastruktur (NFDI)“, Diskussionspapier, April 2017, zuletzt geprüft am 31.01.2018, <http://www.rfii.de/?wpdmdl=2269>; RfII – Rat für Informationsinfrastrukturen, „Zusammenarbeit als Chance: RfII Diskussionspapier März 2018“, zuletzt geprüft am 12.03.2018, <http://www.rfii.de/?wpdmdl=2529>.

kommen wie beschrieben nicht von ungefähr und werden auch nicht von selbst umgesetzt. Vielmehr bedürfen sie eines systematischen Aufbaus von technischen und organisatorischen Infrastrukturen, und zwar von den Datenproduzierenden bis hin zu den Nachnutzenden, ebenso wie des Ausbaus der dazugehörigen Informationskompetenz in allen Bereichen der Hochschulen – und eben längst nicht nur in der Forschung im engeren Sinne. Wir setzen mit unserem Projekt hier an und bauen genau diese notwendige Unterstützung auf, um eine bedeutende Ressource der Zukunft – die digitalen Daten der Wissenschaft – für die Wissenschaft und gleichzeitig auch für die Gesellschaft zu sichern und bestmöglich nutzbar zu machen.

Woher kommen nun die Ressourcen für diese Infrastrukturen an den Hochschulen? Die *Nachhaltigkeit* der an den Hochschulen angesiedelten Infrastrukturen – die sogar der RfII im jüngsten Diskussionspapier⁵ nicht primär in der Verantwortung der NFDI sieht – wird vor allem von den Hochschulen und somit über Landesmittel zu finanzieren sein. Es muss den Ländern klar sein, dass dies nicht unbeträchtliche Mehrinvestitionen bedeutet – auch wenn Bund und andere Förderer sicherlich mit verschiedenen Programmen unterstützen werden, zumindest initial. Die Hochschulen wiederum sind gut beraten, so effizient wie möglich zu agieren, indem sie beispielsweise hochschulübergreifende Kooperationen bilden, um finanziell ökonomisch zu agieren.

Das Projekt HeFDI steht unter Federführung der Philipps-Universität Marburg. Es entstand neben dem oben genannten Positionspapier und dem Kontext der wissenschaftspolitischen Entwicklung auch aus einem landesfinanzierten Pilotprojekt heraus, das die Philipps-Universität Marburg durchgeführt hat: Das damalige *Kompetenzzentrum Forschungsdatenmanagement und -archivierung* hat gemeinsam mit außeruniversitären Einrichtungen wie Archiven und Landesämtern erstmals lokale Strukturen aufgebaut und erste Forschungsprojekte beraten. Schließlich folgte die Koordination aller nun an HeFDI Beteiligten zu einem gemeinsamen Antrag an das Ministerium. Das Ergebnis dieses institutionellen Kooperationsprozesses ist HeFDI: ein Projekt, das Forschungsdateninfrastrukturen an den Hochschulen aufbaut bzw. weiterentwickelt und zugleich vernetzt und Effizienz- sowie Synergieeffekte schafft, damit die Forschenden und andere Interessierte langfristig auf digitale Daten entlang des Datenlebenszyklus zugreifen, diese kombinieren und neu auswerten sowie den Herstellungsweg bestehender Forschungsergebnisse nachvollziehen können.

Diese Ziele werden in HeFDI in einem kooperativen Gesamtkonzept gebündelt und durch inhaltliche Arbeit in sieben Bereichen adressiert:

- *Beratung*: An den beteiligten Hochschulen werden Servicestellen für Forschungsdaten aufgebaut, die Beratungstätigkeiten über den gesamten Datenlebenszyklus anbieten, insbesondere aber bereits im Vorfeld von Forschungsprojekten bzw. der Antragstellung zu Datenstrategien und Datenmanagementplänen. Diese Servicestellen und die dafür eingestellten Forschungsdatenreferent/inn/en sind auf der Arbeitsebene Dreh- und Angelpunkte für die weitere Entwicklung zu Forschungsdatenmanagement an den beteiligten Hochschulen. Sie sind gleichzeitig in eine klare Governancestruktur eingebunden. Die Beratung insbesondere von

5 Ebd.

Drittmittelanträgen läuft aktiv an mindestens fünf Standorten und umfasst nahezu sämtliche fachliche Bereiche. Die Zahl der Beratungsfälle nimmt dabei stetig zu und verdeutlicht den hohen Bedarf der Forschenden an entsprechender Unterstützung bei Datenmanagementplänen und -strategien. Herausfordernde und/oder neuartige Beratungsfälle werden dabei von den Forschungsdatenreferent/inn/en durchaus auf der Arbeitsebene von HeFDI eingebracht und dort per Erfahrungs- und Wissensaustausch beraten.

- *Erarbeitung von Leitlinien:* Den Rahmen für die nachfolgenden Arbeitsbereiche stellen Leitlinien zum Umgang mit Forschungsdaten dar. Jede der Verbundhochschulen hat die Aufgabe, eine solche Leitlinie zu verabschieden, um das Selbstverständnis der jeweiligen Hochschule und ihrer Mitglieder im Umgang mit Forschungsdaten zu formulieren bzw. zu schärfen. Von den Partnerhochschulen haben derzeit bereits sieben eine Leitlinie verabschiedet, bei drei weiteren durchläuft sie gerade den Gremienprozess.⁶
- *Bedarfsanalyse:* Die Hochschulen erfassen kontinuierlich den Bedarf der Forschenden hinsichtlich ihrer digitalen Forschungsdaten. Während einige Universitäten der landesweiten Initiative bereits früh Umfragen abgeschlossen haben,⁷ führen derzeit weitere Hochschulen quantitative und qualitative Bedarfsanalysen bei ihren Forschenden zum Umgang mit digitalen Forschungsdaten durch.
- *Aufbau digitaler Informationskompetenz:* In diesem Arbeitsbereich entwickelt HeFDI sukzessive Schulungsformate und -module zum Umgang mit Forschungsdaten für unterschiedliche Zielgruppen. Zu Beginn stehen dabei modellhaft First-Level-Schulungen, etwa die Vermittlung von Grundlagenkenntnissen zu Metadaten und Datenversionierung. Solche Formate wurden bereits ab Winter 2016 an Pilothochschulen angeboten, werden schrittweise von den kooperierenden Verbundhochschulen übernommen und jeweils an die lokalen Bedarfe und das örtliche Schulungsprogramm angepasst. Darauf aufbauend hat die AG an einigen Standorten fachspezifische Schulungen für disziplinäre Forschungsgruppen entwickelt und/oder organisiert, etwa durch Hinzuziehung von Expert/inn/en. Die enge Kooperation ermöglicht dabei eine zügige Nachnutzung von Schulungsmaterial und Erfahrungen über beide Ausbaustufen und unterschiedliche Veranstaltungsformate hinweg.
- *Repositorien:* Als einen wesentlichen Kern beinhaltet HeFDI die Entwicklung von technischen Lösungen für den Bedarf der Forschenden an Angeboten zur Sicherung, Aufbewahrung und ggf. Veröffentlichung der Forschungsdaten. Die Voraussetzung für gemeinsam entwickelte

- 6 Bereits 2015 hat die Technische Universität Darmstadt ihre Leitlinien zum Umgang mit Forschungsdaten veröffentlicht (Technische Universität Darmstadt, „Leitlinien zum Umgang mit digitalen Forschungsdaten an der TU Darmstadt.“ zuletzt geprüft am 13.04.2018, https://www.tu-darmstadt.de/media/dezernat_vi/relaunch_2015/gute_wiss_praxis/Leitlinien_Forschungsdaten_2015.pdf). Auch die Universität Kassel („Forschungsdaten-Leitlinie: Grundsätze zum Umgang mit Forschungsdaten an der Universität Kassel.“ <https://www.uni-kassel.de/themen/?id=39306>) hat zügig eine Leitlinie aufgestellt. An der Philipps-Universität Marburg („Grundsätze zum Umgang mit Forschungsdaten an der Philipps-Universität Marburg vom 19.12.2017.“ <https://www.uni-marburg.de/de/universitaet/administration/amtliche-mitteilungen/jahrgang-2018/04-2018.pdf>) ist die Leitlinie bereits veröffentlicht. An der Justus-Liebig-Universität Gießen, der Technischen Hochschule Mittelhessen, der Goethe-Universität Frankfurt und der Hochschule Darmstadt sind die Leitlinien jeweils verabschiedet und werden bald publiziert.
- 7 Vgl. Esther Krähwinkel, „Forschungsdatenmanagement an der Philipps-Universität Marburg. Die Ergebnisse der Umfrage zum Forschungsdatenmanagement im November 2014.“ <http://doi.org/10.17192/es2015.0019>; Gerald Langhanke und Wolfgang Stille, „Umgang mit Forschungsdaten – erste Schritte zur Bedarfserhebung und Leitlinienentwicklung: Workshop der hessischen Hochschulen zum Forschungsdatenmanagement.“; Frank Waldschmidt-Dietz und Christian Krippes, „Forschungsdaten an der JLU Gießen: Auswertung einer Umfrage aus dem Juli 2016.“ http://geb.uni-giessen.de/geb/volltexte/2017/12603/pdf/Bericht_Umfrage_FD_2017.pdf.

Infrastrukturlösungen sind dabei gemeinsame Mindeststandards für Metadaten, mit denen die dort abgelegten Daten erschlossen werden.

- *Rechtliche Fragen:* HeFDI baut einerseits Wissen und Kompetenz zu Rechtsfragen auf, andererseits stellt es den Forschenden Informationen zu einschlägigen Rechtsfragen bereit. Letzteres erfolgt in der Regel bedarfsorientiert etwa anhand von konkreten Drittmittelanträgen. Des Weiteren ist es zukünftige Aufgabe des Projekts, eine Gruppe von Expert/inn/en zu Rechtsfragen in Bezug auf Forschungsdatenmanagement ins Leben zu rufen.
- *Aktives Datenmanagement:* Das aktive Datenmanagement, d. h. das Management während der unmittelbaren Arbeit an den Daten, bildet einen eigenen Arbeitsbereich von HeFDI. Diesbezüglich stellen wir den Forschenden zunächst Informationen zu existierenden forschungsunterstützenden Tools zum Handling von Daten zur Verfügung. Dies erfolgt stets priorität orientiert an etablierten und frei verfügbaren sowie innerhalb der Informationsinfrastruktur der Hochschulen integrierten Instrumenten. Dazu zählen etwa die Abfrage und Entwicklung von Workflows oder die Einführung und Vermittlung von Versionierungssystemen wie GIT. Einen großen Schwerpunkt in diesem Bereich stellen elektronische Laborbücher dar, auf die sich die Arbeit aktuell konzentriert.

3. Governance und institutionelle Kooperation

Die Hochschulen in Hessen haben HeFDI über einen längeren Zeitraum und mit vielen Beteiligten entwickelt sowie das Projekt grundlegend in die Hochschullandschaft in Hessen integriert. Im Folgenden stellen wir die dabei kooperativ etablierten Governancestrukturen vor, die aus unserer Sicht für sich genommen ein bereits substanzielles Projektergebnis repräsentieren.

3.1. Projektstruktur

Die Governance nimmt ihren Ausgangspunkt an den beteiligten Hochschulen: Beteiligt sind alle Universitäten und Hochschulen für angewandte Wissenschaften in Hessen (vgl. Abb. 1). Besonders hervorzuheben ist, dass bei diesem Projekt an fast allen Standorten die jeweiligen Hochschulbibliotheken und -rechenzentren entweder unmittelbar beteiligt sind oder sogar als Kernträger des Projekts fungieren. Die Infrastruktureinrichtungen der Hochschulen wirken demnach hochschulübergreifend arbeitsteilig zusammen, um gemeinsam neue Angebote aufzubauen bzw. ihre bestehenden Dienste weiterzuentwickeln. Auch die Forschungsabteilungen sind an einigen Standorten eng eingebunden. Die Zentrale des HeBIS-Verbundes ist ebenfalls in das Projekt involviert und unterstützt es unmittelbar durch Expertise in ihren Kernbereichen. Hier wurde beispielsweise eine Evaluation bestehender Softwarelösungen für Forschungsdatenrepositorien durchgeführt.

Konkrete Umsetzung erfährt das Projekt durch die Einrichtung einer Servicestelle für Forschungsdaten an jeder Institution des Projektverbundes. Für jede beteiligte Hochschule ist dafür im Projektplan eine Stelle für eine/n Forschungsdatenreferenten/in vorgesehen. Eine weitere Stelle mit informationstechnischem Schwerpunkt ist bei HeBIS angesiedelt. Diese Stellen werden flankiert durch solche, die durch die Hochschulen aus eigenem Budget bzw. erfolgreichen Drittmittelanträgen finanziert werden, sodass das Team – mit wachsender Tendenz – auf der Arbeitsebene aktuell 20 Personen umfasst. Dabei ist hervorzuheben, dass ausnahmslos alle beteiligten Hochschulen von Beginn des

Projekts an Kontaktpersonen benannt und Schnittstellen zu ihren Einrichtungen hergestellt haben. So konnten auch zu einem Zeitpunkt, als aufgrund des Stellenbesetzungsprozesses noch keine Forschungsdatenreferent/in vor Ort vorhanden war, Informationen in die jeweilige Verbundhochschule bzw. ihre zuständige Einrichtung fließen. Information und Kommunikation waren so von Beginn des Projekts an gewährleistet.

Das HMWK ist Geldgeber dieses Projekts: HeFDI wird aus dem Innovations- und Strukturentwicklungsbudget des Landes Hessen finanziert, mit der beachtlichen Summe von 3,4 Millionen Euro für etwas mehr als viereinhalb Jahre. Jede Universität ist dabei mit einer vollen, jede Hochschule für angewandte Wissenschaften mit einer halben Stelle vertreten. Zusätzlich wird an mehreren Standorten wird das Projekt durch lokale Mittel aufgestockt, um die lokale Infrastruktur zügig aufzubauen.

3.2. Entscheidungsprozesse

Angesichts des hier umrissenen großen und heterogenen Kreises von Beteiligten besteht eine besondere Herausforderung darin, Entscheidungs-, Kommunikations- und Informationsstrukturen aufzubauen, die ermöglichen, dass alle Beteiligten effizient zusammenarbeiten und nötige Entscheidungen zügig getroffen werden können. HeFDI ist bezüglich der Governance wie folgt aufgestellt (vgl. Abbildung 1):

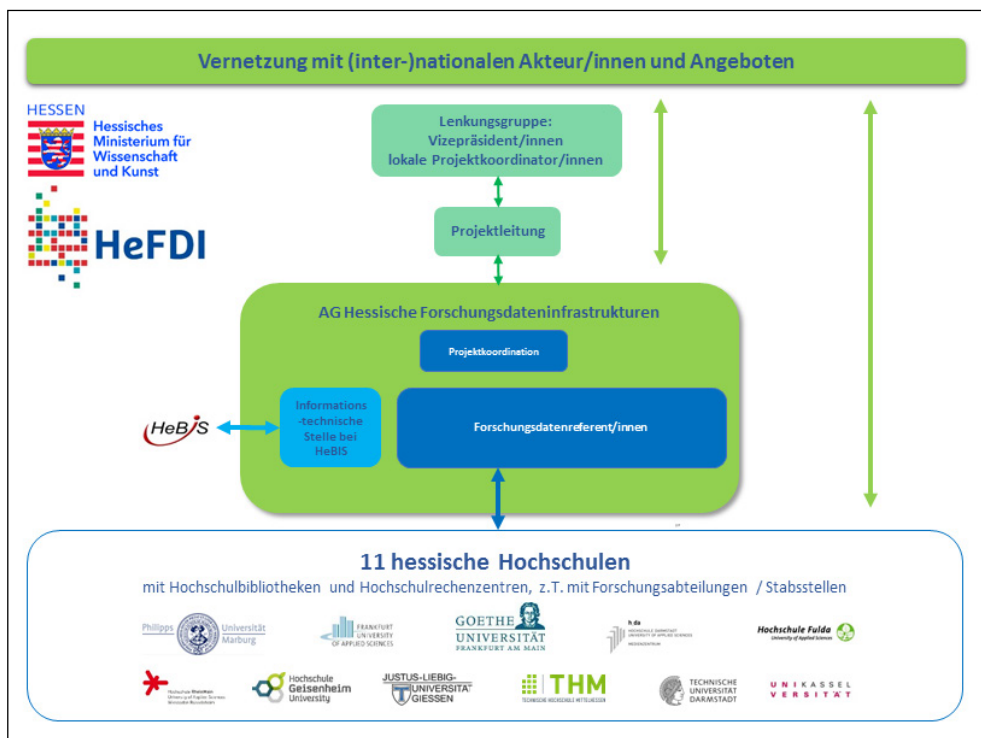


Abb. 1: Governancestruktur des Projekts HeFDI (Quelle: Eigene Darstellung).

Oberste Entscheidungsebene ist die Lenkungsgruppe. Diese besteht aus den für das Thema zuständigen Vizepräsident/inn/en der beteiligten Hochschulen sowie aus den lokalen Projektkoordinator/inn/en. Die Lenkungsgruppe fällt Grundsatzentscheidungen zur Steuerung des Projekts. Die Projektleitung liegt bei Prof. Dr. Joachim Schachtner, Vizepräsident für Informations- und Qualitätsmanagement der Philipps-Universität Marburg. Die Projektkoordination (Dr. Ortrun Brand) und die Geschäftsstelle sind ebenfalls in Marburg angesiedelt. Hier werden alle Arbeitsergebnisse zusammengeführt, sie werden mit der Projektleitung abgestimmt und entsprechende Entscheidungen vorbereitet. Ebenso wird an dieser Stelle die Kommunikation des Gesamtprojekts gestaltet. Projektleitung und Projektkoordination berichten zudem regelmäßig an die Arbeitsgruppe Infrastruktur der Landeshochschulentwicklungsplanung, um auch hier alle Beteiligten sowie das Ministerium als Auftraggeber über den aktuellen Stand sowie Weichenstellungen des Projekts zu informieren.

Die aus den Projektmitteln eingestellten Forschungsdatenreferent/inn/en bilden die Arbeitsebene und damit die AG HeFDI. Diese wird einerseits durch die Geschäftsstelle an der Philipps-Universität Marburg koordiniert, andererseits obliegt ihr ein hohes Maß an Selbstständigkeit durch die thematische Arbeit in Unterarbeitsgruppen.⁸ Deren Inhalte und Struktur sowie die damit verbundene selbstständige Arbeit und Gestaltung erfordern und zugleich ermöglichen.

Zu dieser grundlegenden Governance gesellt sich die Aufgabe der Einbindung jeweiliger lokaler Einrichtungen an den Standorten. Dazu ist zunächst festzuhalten, dass – wie oben beschrieben – die Servicestellen für Forschungsdaten lokal jeweils unterschiedlich angesiedelt sind: an den Bibliotheken, an den Rechenzentren, zum Teil in Hybridkonstruktionen an beiden Einrichtungen, teilweise auch an den Forschungsabteilungen, als Stabsstelle oder direkt in der Präsidialverwaltung der Hochschule. Jede Hochschule hat zudem eine lokale Projektstruktur etabliert, mit der die Arbeitsebene, die lokalen Leitungsebenen der Bibliotheken, Rechenzentren und z.T. der Forschungsabteilungen sowie die lokal zuständigen Vizepräsident/inn/en in das Projekt involviert sind, so dass lokale Kompetenzen zum Forschungsdatenmanagement gebündelt werden. Dadurch ist gewährleistet, dass quer durch die hessische Hochschullandschaft die Entscheidenden, Infrastruktureinrichtungen und weitere einschlägige Akteure – wie etwa die Forschungsabteilungen – systematisch in Information, Entscheidungsprozesse und Ausgestaltung eingebunden sind.

3.3. Institutionelle Arbeitsteilung

Dem Verbundprojekt liegt die Idee einer institutionellen Arbeitsteilung zugrunde. Wie genau diese ausgestaltet wird, ist derzeit noch im Aushandlungsprozess begriffen, und diese zu entwickeln ist Teil und Gegenstand des Projekts. Die Grundintention des Projekts HeFDI richtet sich jedoch auf den kooperativen, arbeitsteiligen Aspekt: Ein Verbundprojekt von und an Hochschulen kann naturgemäß nur dort ansetzen und Infrastrukturen aufbauen, wo die Daten zuerst entstehen und verarbeitet werden. Die beteiligten Hochschulen können sich qua Verbund wechselseitig bei fachspezifischen Schwerpunkten ergänzen, je nach Hochschulkompetenzen und -interessen. Ebenso können die beteiligten Hochschulen als dauerhafte Knoten in einem supranationalen Netzwerk agieren – ein Umstand

⁸ Bislang arbeiten folgende Unterarbeitsgruppen (UAGs): UAG Repositorium, UAG Research Data Management Organisier, UAG Gesamtkonzept, UAG Langzeitarchivierung, UAG Schulung und Beratung.

und ein Bedarf, den bereits Bendix et al. (vgl. auch Bach et al.) formuliert haben.⁹ Abschließend ist zu diskutieren, wie der so dringend nötige Aufbau von Strukturen, Wissen, Kompetenzen in Sachen Forschungsdatenmanagement in die Fläche gebracht und verankert werden kann. An genau diesem Punkt setzt die mit HeFDI intendierte und derzeit ausbuchstabierte institutionelle Arbeitsteilung an und erzielt substanziellen Mehrwert: HeFDI produziert Synergieeffekte durch Bündelung landesweiter Aktivitäten. Dies bringt Spill-over-Effekte mit sich (s.u.), d.h. erst aus der Zusammenarbeit springt einerseits der Funke über und es ergeben sich andererseits weitere Anstöße zu neuen, innovativen Entwicklungen und Services. Die Kooperation ermöglicht zudem einen schnellen Erfahrungsaustausch. HeFDI setzt somit arbeitsteilig an, anstatt parallel Identisches zu entwickeln. Das Projekt kann aufgrund der versammelten fachspezifischen Expertise auch fachspezifische Schwerpunkte bilden, je nach Hochschulkompetenzen und -interessen. HeFDI kann zudem Strukturen, Wissen und Kompetenzen weiträumig verbreiten. Zugleich verfügen wir über ein größeres Verhandlungspotential durch ein entsprechend großes Konsortium – und schließlich wird damit vor allem insgesamt eine Kooperationskultur weiter etabliert, von der alle profitieren.

4. Vernetzung, Spill-over und nächste Schritte

4.1. Vernetzung

HeFDI stellt mit den oben beschriebenen Aspekten ein Projekt dar, dessen Grundstruktur und Prozesse auf ein arbeitsteiliges Vorgehen innerhalb eines Verbundes angelegt sind. Gleichzeitig erhebt dieses Projekt, jenseits des Verbundes, angesichts der eingangs erwähnten Entwicklung den Anspruch auf enge Vernetzung mit nationalen und internationalen Initiativen und Vereinigungen. Die an HeFDI beteiligten Hochschulen engagieren sich deshalb in unterschiedlichen Arbeitsgruppen von DINI/nestor, bei der Research Data Alliance oder sind in Arbeitsgruppen der Allianz-Schwerpunktinitiative „Digitale Information“ präsent. Auch findet ein enger Austausch mit anderen Landesinitiativen sowie einzelnen Forschungsdatenmanagement-Einrichtungen statt. Schließlich werden die Entwicklungen rund um die European Open Science Cloud (EOSC), den EOSC-Pilot und den EOSC-Hub aktiv verfolgt.

4.2. Spill-over

Die mit HeFDI und dem dazugehörigen Verbund gebildete Kooperation hat binnen kürzester Zeit erhebliche Erfolge im Sinne von Spill-over-Effekten gezeitigt: Bereit Ende 2016/Anfang 2017 konnte mit fünf der an HeFDI beteiligten Hochschulen das BMBF-Projekt „Forschungsdatenkurse für Studierende und Graduierte (FOKUS)“¹⁰ eingeworben werden. An zwei Standorten haben Sonderforschungsbereiche mit expliziter Unterstützung der HeFDI-finanzierten Servicestellen für Forschungsdaten erfolgreich Informations-Infrastrukturprojekte für Forschungsdatenmanagement eingeworben.¹¹

9 Jörg Bendix, Jens Nieschulze und William K. Michener, „Data platforms in integrative biodiversity research,“ *Ecological Informatics*, Nr. 11 (2012), <http://doi.org/10.1016/j.ecoinf.2012.04.001>; Kerstin Bach et al., „A comparative evaluation of technical solutions for long-term data repositories in integrative biodiversity research,“ *Ecological Informatics*, Nr. 11 (2012), <http://doi.org/10.1016/j.ecoinf.2011.11.008>.

10 Vgl. „FOKUS – Forschungsdatenkurse für Studierende und Graduierte,“ Philipps-Universität Marburg, zuletzt geprüft am 21.06.2018, <https://www.uni-marburg.de/projekte/forschungsdaten/projekt/fokus>.

11 Vgl. „Informationsinfrastruktur,“ Technische Universität Darmstadt, zuletzt geprüft am 13.04.2018, https://www.sfb1194.tu-darmstadt.de/teilprojekte_4/z_inf/z_inf.de.jsp sowie „NeuroScientific Workflow Assistance,“ Justus-Liebig-Universität Gießen/Philipps-Universität Marburg, zuletzt geprüft am 13.04.2018, <http://www.allpsych.uni-giessen.de/sfb/projekte/inf.html>.

Anfang 2018 hat HeFDI beim BMBF-Ideenwettbewerb „Wissenschaft im digitalen Wandel“ mit den Projekten „Training zum sensiblen Umgang mit Forschungsdaten (TRUST)“ (Philipps-Universität Marburg) und „FDM-Planung im Maschinenbau“ (TU Darmstadt/RWTH Aachen) zwei der fünf zu vergebenen Preise gewonnen. An zwei Standorten werden darüber hinaus die Aktivitäten rund um die Formierung von NFDI-Konsortien vorangetrieben. Diese Beispiele zeigen, dass der schnelle Wissensaufbau und -austausch sowie die stabile Verbundstruktur rasant Früchte tragen.

4.3. Nächste Schritte

Die nächsten Schritte des HeFDI-Projektes bestehen in der Umsetzung technischer Lösungen zum Management von Forschungsdaten in Form von institutionellen Repositorien. Entsprechende Vorarbeiten sind bereits abgeschlossen. Geplant sind vorerst zwei Pilotrepositorien in Marburg und Darmstadt, wobei zum einen – in Marburg – die gemeinsam im Projekt erarbeiteten Use-Cases für ein Forschungsdatenrepositorium abgebildet werden und dabei eine Repositoriums-Basisinstallation und -konfiguration entwickelt wird. Zum anderen werden – in Darmstadt – zusätzlich Workflows rund um die Anbindung eines Repositoriums an Anwendungen im Big Data-Bereich abgebildet. Es stehen zudem Überlegungen im Raum, aus lokalen Schwerpunkten heraus auf Basis der beiden Pilotprojekte fachliche Repositoriums-Angebote aufzubauen, etwa für Daten aus den Geisteswissenschaften oder für Daten aus den Ingenieurs- und Naturwissenschaften. So können arbeitsteilig auch disziplinspezifische Ansätze verfolgt werden, die spezielles Knowhow benötigen und daher im rein institutionellen Kontext aus Ressourcengründen oftmals schwer oder gar nicht umsetzbar sind. Vorstellbar wäre unter anderem, die Services mittelfristig auch Datenproduzierenden anderer Hochschulen in Hessen und darüber hinaus anzubieten. So könnte disziplinspezifisches Datenmanagement im Verbund aufgebaut und institutionell betrieben und auch finanziert werden. Diese Konstruktion war bisher fast ausschließlich außeruniversitären Forschungseinrichtungen vorbehalten und stellt in der aktuellen Konstellation in Hessen ein Novum dar. Die Repositoriums-Lösungen der hessischen Kooperation werden zudem über entsprechende Schnittstellen und Metadaten-Standards aktiv mit größeren Angeboten vernetzt. Zudem entwickelt HeFDI ein auf den bisherigen Praxiserfahrungen und Pilotierungen basiertes Gesamtkonzept, wozu die DFG eine Stellungnahme abgeben wird. In der Gesamtschau wird in Kombination mit weiteren bereits in der Umsetzung befindlichen Maßnahmen, etwa dem flächendeckenden Einsatz von Research Data Management Organiser zur Planung von Forschungsdatenmanagement, aus HeFDI ein zukunftsfähiges und nachhaltiges Gesamtpaket, welches bestens gerüstet ist für kommende nationale Aktivitäten.

5. Fazit

Das Projekt HeFDI ist im Rahmen einer langfristigen Entwicklung entstanden, die die wissenschaftspolitischen Strukturen des Landes Hessen sowie alle beteiligten Akteure eingebunden hat. Diese systematische Entwicklung in Verbindung mit der Governancestruktur des Projekts lässt sich auch als Middle-out-Ansatz kennzeichnen: HeFDI ist weder top-down vom Projektauftraggeber gesteuert noch ausschließlich bottom-up von der Arbeitsebene oder aus den Bibliotheken bzw. Rechenzentren heraus entwickelt, sondern bündelt Impulse von und auf allen Ebenen und realisiert stabile Entscheidungen Middle-out, das heißt unter einer grundlegenden Einbeziehung der Entscheidungsebene

an den Hochschulen. Gleichzeitig ist es eingebunden in eine Gesamtstrategie der Hochschulen des Landes Hessen.

HeFDI ist des Weiteren gekennzeichnet durch sieben klar definierte Arbeitsbereiche. Alle relevanten Akteure der hessischen Hochschulen sind eingebunden. Prägend für das Projekt sind Beteiligung, Kooperation und Arbeitsteilung sowie die strategische Bündelung der Aktivitäten durch Koordination. Bereits jetzt zeitigt diese Struktur viele positive Effekte.

Schließlich setzt HeFDI als Projekt dort an, wo eine fundierte Architektur und Infrastruktur für Forschungsdaten entsteht bzw. bereits vorhanden ist: an den Hochschulen, wo die Daten initial produziert werden. Die Hochschulen haben erkannt, dass ein Zusammenschluss zum Verbund sinnvoll ist, um ihre Kompetenzen und Angebote im Feld auszutauschen und zu ergänzen. Das kooperative Vorgehen wird somit zur tragenden Grundlage, weil durch den Austausch und die Bündelung von Aktivitäten und Kompetenzen effizientes Handeln möglich ist. Gleichzeitig wird durch den engen Dialog mit den Forschenden und den mit ihnen gemeinsam entwickelten Infokompetenzangeboten dezentrale Vernetzung gewährleistet.

Forschungsdatenmanagement braucht ein strukturelles und vor allem tragendes Netzwerk, welches dort ansetzt, wo die Daten entstehen. Forschung und die von ihr produzierten Daten kennen keine bzw. kaum noch Grenzen. Umso wichtiger ist es, von Beginn des Forschungsprozesses an oder gar noch vorher, bei dessen Planung, Unterstützung für ein professionelles und nachhaltiges Datenmanagement zu leisten. Das Netzwerk wird erst dann vollständig in Bezug auf den Forschungsdatenlebenszyklus, wenn Forschende bereits bei der Entstehung von Daten oder deren Planung Unterstützung erfahren.

Föderale sowie disziplinäre Verbundprojekte stärken dieses Netzwerk zusätzlich. Dadurch entstehen Synergien und die Vernetzung mit Akteuren außerhalb Hessens. Diese machen HeFDI anschlussfähig an nationale Entwicklungen rund um die NFDI. Gleichzeitig werden Universitäten und Hochschulen für angewandte Wissenschaften durch die Bundesländer finanziert. Aufgrund dieser Tatsache und aufgrund der zahlreichen lokalen Facetten des Datenmanagements ist anzustreben, dass Zukunftsthemen wie Forschungsdateninfrastrukturen und das dazugehörige Forschungsdatenmanagement auch von Seiten der Länder dauerhaft und nachhaltig mit Ressourcen ausgestattet werden.

Literaturverzeichnis

- Allianz der Deutschen Wissenschaftsorganisationen. „Grundsätze zum Umgang mit Forschungsdaten.“ Berlin, 2010. Zuletzt geprüft am 25.05.2018. http://www.wissenschaftsrat.de/download/archiv/Allianz_Grundsaeetze_Forschungsdaten.pdf.
- Bach, Kerstin, Daniel Schäfer, Neela Enke, Bernhard Seeger, Birgit Gemeinholzer und Jörg Bendix. „A comparative evaluation of technical solutions for long-term data repositories in integrative biodiversity research.“ *Ecological Informatics*, Nr. 11 (2012): 16–24. <https://doi.org/10.1016/j.ecoinf.2011.11.008>.

- Bendix, Jörg, Jens Nieschulze und William K. Michener. „Data platforms in integrative biodiversity research.“ *Ecological Informatics*, Nr. 11 (2012): 1–4. 10.1016/j.ecoinf.2012.04.001.
- DFG – Deutsche Forschungsgemeinschaft. „Leitlinien zum Umgang mit Forschungsdaten.“ Bonn, 2015. Zuletzt geprüft am 25.05.2018. http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf.
- HRK – Hochschulrektorenkonferenz. „Hochschule im digitalen Zeitalter: Informationskompetenz neu begreifen – Prozesse anders steuern; Entschließung der 13. Mitgliederversammlung der HRK am 20. November 2012 in Göttingen.“ Bonn, 2012.
- HRK – Hochschulrektorenkonferenz. „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können: Orientierungspfade, Handlungsoptionen, Szenarien; Empfehlung der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel.“ Bonn, 2015. Zuletzt geprüft am 25.05.2018. <https://www.hrk.de/positionen/beschluss/detail/wie-hochschulleitungen-die-entwicklung-des-forschungsdatenmanagements-steuern-koennen-orientierungsp/>
- Krähwinkel, Esther. „Forschungsdatenmanagement an der Philipps-Universität Marburg. Die Ergebnisse der Umfrage zum Forschungsdatenmanagement im November 2014.“ Marburg, 2016. <http://doi.org/10.17192/es2015.0019>.
- Langhanke, Gerald und Wolfgang Stille. „Umgang mit Forschungsdaten – erste Schritte zur Bedarfserhebung und Leitlinienentwicklung: Workshop der hessischen Hochschulen zum Forschungsdatenmanagement.“ Marburg, 2015.
- LERU League of European Research Universities. „Roadmap for Research Data.“ Leuven, 2013. Zuletzt geprüft am 14.06.2018. <https://www.leru.org/publications/leru-roadmap-for-research-data>.
- RfII – Rat für Informationsinfrastrukturen. „Zusammenarbeit als Chance: RfII Diskussionspapier März 2018.“ Göttingen, 2018. Zuletzt geprüft am 25.05.2018. <http://www.rfii.de/?wpdmdl=2529>.
- RfII – Rat für Informationsinfrastrukturen. „Leistung aus Vielfalt: Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsmanagements in Deutschland.“ Göttingen, 2016. Zuletzt geprüft am 25.05.2018. <http://www.rfii.de/?wpdmdl=1998>.
- RfII – Rat für Informationsinfrastrukturen. „Schritt für Schritt – oder: Was bringt wer mit? Ein Diskussionsimpuls zu Zielstellung und Voraussetzungen für den Einstieg in die Nationale Forschungsdateninfrastruktur (NFDI).“ Diskussionspapier, April 2017. Göttingen, 2017. Zuletzt geprüft am 25.05.2018. <http://www.rfii.de/?wpdmdl=2269> .

- Technische Universität Darmstadt. „Leitlinien zum Umgang mit digitalen Forschungsdaten an der TU Darmstadt.“ Darmstadt, 2015. Zuletzt geprüft am 14.06.2018. https://www.tu-darmstadt.de/media/dezernat_vi/relaunch_2015/gute_wiss_praxis/Leitlinien_Forschungsdaten_2015.pdf.
- Philipps-Universität Marburg. „Grundsätze zum Umgang mit Forschungsdaten an der Philipps-Universität Marburg vom 19.12.2017.“ Marburg, 2018. Zuletzt geprüft am 25.05.2018. <https://www.uni-marburg.de/de/universitaet/administration/amtliche-mitteilungen/jahrgang-2018/04-2018.pdf>.
- Universität Kassel. „Forschungsdaten-Leitlinie: Grundsätze zum Umgang mit Forschungsdaten an der Universität Kassel.“ Kassel, 2017. Zuletzt geprüft am 25.05.2018. <https://www.uni-kassel.de/themen/?id=39306> .
- Waldschmidt-Dietz, Frank und Christian Krippes. „Forschungsdaten an der JLU Gießen: Auswertung einer Umfrage aus dem Juli 2016.“ Gießen, 2017. Zuletzt geprüft am 25.05.2018. http://geb.uni-giessen.de/geb/volltexte/2017/12603/pdf/Bericht_Umfrage_FD_2017.pdf .
- Wissenschaftsrat. „Empfehlungen zu wissenschaftlicher Integrität. Positionspapier.“ Stuttgart, 2015. Zuletzt geprüft am 25.05.2018. <http://www.wissenschaftsrat.de/download/archiv/4609-15.pdf>.

Von der Idee zum Konzept – Forschungsdatenmanagement an der Universität zu Köln

Jens Dierkes, Universitäts- und Stadtbibliothek Köln

Constanze Curdt, Regionales Rechenzentrum der Universität zu Köln

Zusammenfassung:

Die Universität zu Köln, als eine der größten Hochschulen Deutschlands, nähert sich dem Thema universitätsweites systematisches Forschungsdatenmanagement (FDM) über eine Machbarkeitsstudie an. Im Laufe eines Jahres wurde der Status quo des Umgangs mit Forschungsdaten an der Universität, den Fakultäten, Instituten und Forschungsprojekten ermittelt. Als Grundlage für die weiteren Arbeiten wurde eine Leitlinie zum Umgang mit Forschungsdaten erarbeitet und seitens der Universität verabschiedet. Ausgehend von einem umfangreichen FDM-Service-Portfolio wurden erste Maßnahmenpakete entwickelt, die mit einer realistischen Aufwandsabschätzung eine Grundlage für ein universitätsweites FDM innerhalb der nächsten drei Jahre legen sollen. Die Maßnahmen basieren im Wesentlichen auf dem Aufbau von Informations-, Beratungs- und Schulungsangeboten und sollen die Vernetzung der FDM-Akteure stärken. Ein weiteres Arbeitsgebiet liegt im Aufbau digitaler Services im Bereich Speicherung und Sichtbarmachung von Forschungsergebnissen.

Summary:

The University of Cologne, one of the largest universities in Germany, has approached the topic of university-wide, systematic research data management (RDM) by means of a feasibility study. In the course of a year, the status quo of the handling of research data at the university, the faculties, institutes and research projects was investigated. As a basis for further work, a guideline for the handling of research data was developed and adopted by the university. Based on a comprehensive RDM-service portfolio, first packages of measures were developed in order to provide a basis for a university-wide RDM within the next three years, also giving a realistic estimate of costs. Essentially, the measures are centred on the development of information, consulting and training services and are intended to strengthen the networking of RDM actors. Another field of activity is the development of digital services in the area of storage and visualisation of research results.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S28-46>

Autorenidentifikation: Dierkes, Jens: GND 1067541527, ORCID <http://orcid.org/0000-0002-0121-9261>; Curdt, Constanze: ORCID <http://orcid.org/0000-0002-9606-9883>

Schlagwörter: Forschungsdatenmanagement; Institutionelle Lösung; Informationsinfrastruktur

1. Forschungsdatenmanagement als institutionelle Aufgabe

Der Umgang mit Forschungsdaten, die im Rahmen von wissenschaftlichen Vorhaben entstehen, hat in den letzten Jahren sehr an Bedeutung gewonnen und wurde als Thema und Aufgabe erkannt. Entsprechend wurden diverse Empfehlungen und Positionspapiere zum Thema Forschungsdatenmanagement

(FDM) veröffentlicht.¹ Seitens einiger Forschungsförderer gibt es die Anforderungen, Forschungsdaten nachhaltig aufzubewahren, zugänglich und nachnutzbar zu machen. Diese werden in den jeweiligen Richtlinien der Forschungsförderer und Förderprogramme (z.B. Deutsche Forschungsgemeinschaft DFG, Bundesministerium für Bildung und Forschung BMBF, 8. Rahmenprogramm der Europäischen Union Horizon 2020) beschrieben. Auch erfordern die zunehmend komplexeren Forschungsfragen ein stärkeres Maß an systematischem Management von Forschungsdaten, insbesondere in inter- bzw. transdisziplinären Forschungsvorhaben. Dazu gehört auf der einen Seite die Vernetzung von Forschenden und die damit verbundene Mobilität von Daten (Team Science)² und auf der anderen Seite die mit der Integration von Daten aus verschiedenen Quellen und unterschiedlichen Kontexten verbundenen Herausforderungen³.

Auf internationaler und nationaler Ebene ist das FDM seit vielen Jahren etabliert⁴. Verschiedene Projekte und Initiativen haben sich mit der Thematik in den einzelnen Fachdisziplinen und z.T. fächerübergreifend beschäftigt und Lösungen entwickelt. Hochschulen haben Forschungsdaten und die zugrundeliegende Informationsinfrastruktur als strategische Herausforderung erkannt.⁵ Auf Hochschulebene wurden an vielen Standorten Policies für den Umgang mit Forschungsdaten erlassen, entsprechende Strategien festgelegt und umgesetzt (z. B. Beratungsangebot, Schulungsangebot, Informations- und Kommunikations-Infrastruktur)⁶.

Für eine Hochschule stellt sich dabei die Frage, wie sie unter den lokalen Bedingungen den Forschenden (und Studierenden) die bestmögliche und gleichzeitig auch nachhaltige digitale Infrastruktur zur Verfügung stellen kann. Dabei ist es eine komplexe Aufgabe, die nötige Balance zwischen

- 1 Vgl. „Management von Forschungsdaten–eine zentrale strategische Herausforderung für Hochschulleitungen. Empfehlung der 16. Mitgliederversammlung der HRK am 13. Mai 2014 in Frankfurt am Main,“ Hochschulrektorenkonferenz HRK, zuletzt geprüft am 23.05.2018, <https://www.hrk.de/positionen/beschluss/detail/management-von-forschungsdaten-eine-zentrale-strategische-herausforderung-fuer-hochschulleitungen/>; „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können: Orientierungspfade, Handlungsoptionen, Szenarien. Empfehlung der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel,“ Hochschulrektorenkonferenz HRK, zuletzt geprüft am 23.05.2018, <https://www.hrk.de/positionen/beschluss/detail/wie-hochschulleitungen-die-entwicklung-des-forschungsdatenmanagements-steuern-koennen-orientierungsp/>; „Umgang mit Forschungsdaten,“ DFG, zuletzt geprüft am 27.06.2018, http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/index.html; „Grundsätze zum Umgang mit Forschungsdaten,“ Allianz-Initiative, zuletzt geprüft am 05.03.2018, <https://www.allianzinitiative.de/archiv/forschungsdaten/grundsaeetze>; Rat für Informationsinfrastrukturen, *Leistung aus Vielfalt: Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland*, Göttingen, 2016, <http://nbn-resolving.de/urn:nbn:de:101:1-201606229098>; „Open Access and Data Management,“ EC Horizon 2020, zuletzt geprüft am 05.03.2018, http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm.
- 2 Vgl. Manisha Desai, „The integration of the data scientist into the team: Implications and challenges,“ *Data Science* 1, Nr. 12 (1. Januar 2017): 95100, <https://doi.org/10.3233/DS-170008>.
- 3 Michael Gertz, Carlos Rueda und Jianting Zhang, „Interoperability and Data Integration in the Geosciences,“ in *Scientific Data Management: Challenges, Technology, and Deployment*, hrsg. Arie Shoshani und Doron Rotem, Computational Science Series (CRC Press/Chapman & Hall, 2009), 369–398. Benjamin Kormeier, Klaus Hippe und Ralf Hofestädt, „Data Warehouses in Bioinformatics: Integration of Molecular Biological Data,“ *it - Information Technology Methoden und innovative Anwendungen der Informatik und Informationstechnik* 53, Nr. 5 (2011): 241–249, <https://doi.org/10.1524/itit.2011.0649>.
- 4 Graham Pryor, Sarah Jones und Angus Whyte, *Delivering research data management services: Fundamentals of good practice* (Facet Publishing, 2013). Joyce M. Ray, *Research data management: Practical strategies for information professionals* (Purdue University Press, 2014).
- 5 „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können,“ HRK.
- 6 „Data Policies,“ zuletzt geprüft am 05.03.2018, http://www.forschungsdaten.org/index.php/Data_Policies.

möglichst weiter Verbreitung von Services und fachspezifischer Tiefe zu finden und immer wieder nachzujustieren.⁷

Der DINI/nestor Workshop „Kooperationstreffen Forschungsdaten“ im Oktober 2016 in Göttingen, bei dem Praktikerinnen und Praktiker aus ganz Deutschland zusammengekommen sind, hat gezeigt, dass eine bessere Zusammenarbeit zwischen den FDM-Initiativen und Standorten wünschenswert ist.⁸ Das Thema wurde von den E-Science-Tagen „Forschungsdaten managen“ im März 2017 in Heidelberg wieder aufgegriffen.⁹ Dies könnte zu einer Konsolidierung beitragen und ermöglichen, dass einzelne Standorte sich auf ihre Spezifika und Stärken konzentrieren können, wie beispielsweise im US-amerikanischen Data Curation Network.¹⁰ Diese Vorgehensweise ist auch im Sinne der Empfehlungen des Rates für Informationsinfrastrukturen.¹¹

1.1. Machbarkeitsstudie für ein universitätsweites Forschungsdatenmanagement

An der Universität zu Köln (UzK) erfolgte bisher keine universitätsweite, systematische Auseinandersetzung mit dem Thema FDM. Aktuell werden an der UzK häufig dezentrale Lösungen umgesetzt (auf Projekt-, Instituts-, Fakultätsebene), inklusive der Herausforderung, dass in der Regel die Nachhaltigkeit nicht gesichert ist. Ein Handlungsbedarf, der die gesamte Universität in den Blick nimmt, wird allerdings offensichtlich, wenn man das Vernetzungs- bzw. Synergiepotenzial zusammen mit den diversen Empfehlungen und Richtlinien für den Umgang mit Forschungsdaten, also insbesondere die Antragsfähigkeit (z.B. Exzellenz-Strategie) betrachtet.

Zu diesem Zweck wurde Ende 2016 vom Rektorat der Universität zu Köln eine einjährige Machbarkeitsstudie zum FDM an der UzK in Auftrag gegeben. In diesem Projekt haben sich das Dezernat Forschungsmanagement der Universitätsverwaltung (D7), das Regionale Rechenzentrum (RRZK) und die Universitäts- und Stadtbibliothek (USB) der Universität zu Köln zusammengefunden. Ziel war es, den Status quo und ein Konzept für die Entwicklung eines universitätsweiten FDMs zu ermitteln. Als Lenkungsgremium wurde die Prorektorin für Forschung, die Vize-Kanzlerin, die Forschungs-Prodekaninnen und Prodekane der sechs Fakultäten, sowie die Leitungen der drei beteiligten Einrichtungen eingesetzt. Das Projektteam umfasste im Durchschnitt sieben Mitarbeiterinnen und Mitarbeiter der Projektpartner, die Anteile ihrer Arbeitskraft in die Umsetzung der Studie investierten.

Im Folgenden sollen die zentralen Ergebnisse dieser Machbarkeitsstudie vorgestellt werden. Das Dokument ist wie folgt aufgebaut. In Abschnitt 2 wird kurz auf die Situation des systematischen FDMs in Deutschland eingegangen, insbesondere auf den Status der aktuellen Exzellenz-Universitäten.

7 Vgl. David F. Flanders, „A New Model for Research IT Services“, *David F. Flanders (Blog)*, 15. April 2013, zuletzt geprüft am 05.03.2018, <https://dfflanders.wordpress.com/2013/04/15/resbaz/>.

8 „6. DINI/nestor-Workshop Kooperationstreffen Forschungsdaten,“ zuletzt geprüft am 05.03.2018, <http://www.forschungsdaten.org/index.php/DINI-nestor-WS6>.

9 Jonas Kratzke und Vincent Heuveline, Hrsg., *E-Science-Tage 2017: Forschungsdaten managen* (Heidelberg: heiBOOKS, 2017), <https://doi.org/10.11588/heibooks.285.377>.

10 Lisa Johnston u. a., „Data Curation Network: How Do We Compare? A Snapshot of Six Academic Library Institutions' Data Repository and Curation Services,“ *Journal of eScience Librarianship* 6, Nr. 1 (28. Februar 2017), <https://doi.org/10.7191/jeslib.2017.1102>.

11 Rat für Informationsinfrastrukturen, *Leistung aus Vielfalt*.

Danach wird im Abschnitt 3 der Blick auf die Situation an der UzK gelenkt und die Erarbeitung eines Konzeptes für ein universitätsweites FDM präsentiert. Anschließend werden konkrete Maßnahmen abgeleitet und Ideen für die organisatorische Einbettung vorgestellt.

2. Forschungsdatenmanagement an deutschen Universitäten

Hochschulen, national und international, stehen derzeit vor der Herausforderung, ein professionelles FDM aufzubauen und zu etablieren. Die Unterstützung der Wissenschaftlerinnen und Wissenschaftler im Umgang mit den erhobenen Forschungsdaten steht dabei im Fokus. Gleichzeitig wird durch den Aufbau eines hochschulweiten FDM die Attraktivität und Konkurrenzfähigkeit der Hochschule erhalten und gesteigert.

An einigen nationalen und internationalen Hochschulen wurde diese Aufgabe bereits aufgegriffen und umgesetzt.¹² Entsprechend der Empfehlungen der Hochschulrektorenkonferenz (HRK)¹³ wurden in den letzten Jahren an einzelnen deutschen Hochschulen beispielsweise Leitlinien für den Umgang mit Forschungsdaten aufgestellt, sowie entsprechende Strategien festgelegt und umgesetzt (z.B. Aufbau von Beratungs- und Schulungsangebot, technische Infrastruktur). Einen Überblick über vorhandene FDM-Strukturen, beteiligte Hochschuleinrichtungen und FDM-Service-Angebote an derzeitigen deutschen Exzellenz-Universitäten wird in Tabelle 1 dargestellt. Es ist ersichtlich, dass bei der Umsetzung des FDM an den verschiedenen Standorten teilweise unterschiedliche Ansätze gewählt wurden, beispielsweise im Hinblick auf die Auswahl der beteiligten Institutionen am zentralen FDM. Weiterhin wurde die Veröffentlichung einer Leitlinie nicht bei allen Hochschulen priorisiert. Das FDM Service-Angebot wird an den verschiedenen Standorten sehr unterschiedlich umgesetzt. Als Minimal-Service ist eine FDM-Beratung an allen Standorten etabliert.

Für den Aufbau und die Etablierung von hochschulweiten FDM-Strukturen ist der Austausch und die Kooperation mit bestehenden Netzwerken und Initiativen (beispielsweise DINI/nestor oder RDA), sowie mit anderen Hochschulen essentiell. Gesammelte Erfahrungen von anderen Standorten (u.a. bei der Erstellung einer Leitlinie oder dem Aufbau von FDM-Services) können häufig auf andere Hochschulen übertragen werden. Bei der Konzeption eines FDM an der UzK ist daher auch ein Austausch mit anderen Referenz-Hochschulen und Netzwerken im Rahmen der Machbarkeitsstudie erfolgt.

12 „Forschungsdatenmanagement.“ RWTH Aachen, zuletzt geprüft am 27.02.2018, <http://www.rwth-aachen.de/cms/root/Forschung/~Inaw/Forschungsdatenmanagement/>, vgl. Thomas Eifert, Stephan Muckel und Dominik Schmitz, „Introducing research data management as a service suite at RWTH Aachen university,“ in 9. DFN-Forum Kommunikations-technologien, Bd. P-257, Lecture Notes in Informatics (Gesellschaft für Informatik, Bonn, 2016), 55–64, zuletzt geprüft am 23.05.2018, <http://subs.emis.de/LNI/Proceedings/Proceedings257/article10.html>; Jens Dierkes und Ulrike Wuttke, „The Göttingen eResearch Alliance: A Case Study of Developing and Establishing Institutional Support for Research Data Management“, *ISPRS International Journal of Geo-Information* 5, Nr. 8 (1. August 2016): 133, <https://doi.org/10.3390/ijgi5080133>; Maxi Kindling, Peter Schirmbacher und Elena Simukovic, „Forschungsdatenmanagement an Hochschulen: Das Beispiel der Humboldt-Universität zu Berlin,“ *LIBREAS. Library Ideas*, Nr. 23 (2013), <http://doi.org/10.18452/9041>; „Research Data Management,“ TU Delft, zuletzt geprüft am 27.02.2018, <https://www.tudelft.nl/en/library/current-topics/research-data-management/>.

13 „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können,“ HRK.

Tabelle 1: Übersicht FDM-Aktivitäten an deutschen Exzellenz-Hochschulen (Stand 20.02.2018)

Universität (Status)	Leitlinien-verabschiedung	Beteiligte Institutionen	FDM-Angebote	FDM-Kontakt/ in Aufbau
FU Berlin (U15)		Universitätsbibliothek, Zentraleinrichtung für Datenverarbeitung, Center für Digitale Systeme	im Aufbau (Repositorium, Beratung, Schulung)	Kontaktstelle seit 2015
HU Berlin (U15)	2014	Zentraleinrichtungen Computer- und Medienservice, Servicezentrum Forschung, Universitätsbibliothek und Vizepräsident Forschung	Beratung, Repositorien, Schulungen	Kontaktstelle seit 2012
RWTH Aachen (TU9)	2016	Universitätsbibliothek, IT Center und Zentrale Hochschulverwaltung (Dezernate 1, 4, 5)	Beratung, Weiterbildungsangebot, institutionelles Repositorium u. Archiv,	Service Desk seit 2015
TU Dresden (TU9)	2014 (Teil der Sicherung guter wiss. Praxis)	Sächsische Landesbibliothek Staats- und Universitätsbibliothek, Zentrum für Informationsdienste und Hochleistungsrechnen	Beratung zum Management von Forschungsdaten, Training, Zwischen- und Langzeitarchiv	Kontaktstelle seit 2016
Universität Heidelberg (U15)	2014	Service-Einrichtung des Universitätsrechenzentrums und der Universitätsbibliothek Heidelberg	Beratung, Weiterbildung, Repositorium, Datenbearbeitung, Archivierung und Veröffentlichung	Kompetenzzentrum seit 2013
Universität Konstanz		Kommunikations-, Informations-, Medienzentrum	Beratung, BwFDM-Dienste	Kontaktstelle
TU München (TU9)		Universitätsbibliothek	Beratung, Schulung, institutionelles Repositorium	Beratungszentrum
LMU München		Universitätsbibliothek	Forschungsdaten-repositorium	Kontaktstelle
Universität Bremen		Forschungsförderung	Beratung	Kontaktstelle
Universität Tübingen (U15)	ja	Informations-, Kommunikations- und Medienzentrum, eScience-Center	Beratung, Forschungsdatenarchiv	Kontaktstelle

3. Forschungsdatenmanagement an der Universität zu Köln

3.1. Status quo

An der UzK haben sich in den vergangenen Jahren im Bereich des FDM bereits dezentrale Teillösungen und Kompetenzen gebildet. Gespräche der Projektmitarbeiterinnen und Projektmitarbeiter auf Fakultätsebene und mit Forschungsverbänden haben gezeigt, dass es sich dabei einerseits um diverse Lösungen handelt, die im Rahmen von Forschungsprojekten entstanden sind. Andererseits wurden aber vereinzelt auch breiter angesetzte strukturbildende Maßnahmen auf Instituts- oder Fakultätsebene etabliert.

Auf Fakultätsebene hat die Philosophische Fakultät im Bereich des Aufbaus von zentralen FDM-Strukturen eine Vorreiterrolle innerhalb der UzK eingenommen. Das Data Center for the Humanities (DCH)¹⁴ und das Cologne Center for eHumanities (CCeH)¹⁵ sind etablierte Einrichtungen an der Philosophischen Fakultät. Das Leistungsangebot richtet sich dabei hauptsächlich an alle Forschenden der Fakultät. Das DCH nimmt dabei insbesondere beratende Aufgaben wahr (u.a. Antragsberatung) und versucht dadurch möglichst früh, die Projekte im datenverarbeitenden Bereich zu unterstützen. An den anderen fünf Fakultäten der UzK (Wirtschafts- und Sozialwissenschaften, Rechtswissenschaften, Medizin, Mathematik/Naturwissenschaften, Humanwissenschaften) sind derzeit keine vergleichbaren, zentralen FDM-Strukturen zur Unterstützung der Forschenden der gesamten Fakultät vorhanden.

Bislang wurden an zwei Fakultäten der UzK Umfragen zum Umgang mit Forschungsdaten durchgeführt. An der Philosophischen Fakultät wurde 2016 in Rahmen einer Masterarbeit in Zusammenarbeit mit dem DCH und der Universitäts- und Stadtbibliothek Köln (USB) eine Umfrage unter den Forschenden zum Thema Forschungsdaten durchgeführt. Die Ergebnisse der Studie sind in dem Bericht „Forschungsdaten an der Philosophischen Fakultät der Universität zu Köln“¹⁶ zusammengefasst. Eine vergleichbare Studie wurde an der Wirtschafts- und Sozialwissenschaftlichen Fakultät 2017 im Rahmen einer Projektarbeit seitens der USB durchgeführt.¹⁷

An einigen Fakultäten der UzK sind über die letzten Jahre (teilweise sogar Jahrzehnte) diverse FDM-„Insellösungen“ auf Instituts- oder Arbeitsgruppenebene gewachsen. Größtenteils handelt es sich dabei um fachspezifische, technische Infrastrukturen zur Speicherung und Archivierung von Forschungsdaten. Etablierte Strukturen gibt es beispielsweise in der Medizinischen Fakultät, der Mathematisch-Naturwissenschaftlichen Fakultät (z.B. Institut für Geophysik und Meteorologie, Physikalische Institute) oder der Philosophischen Fakultät (z.B. Institut für Ur- und Frühgeschichte, Archäologisches Institut).

Auf Projektebene gibt es diverse FDM-Aktivitäten an der UzK. An dieser Stelle werden nur einzelne, ausgewählte Initiativen genannt. Im Rahmen von größeren Verbundprojekten ist es seit einigen Jahren

14 Data Centre for the Humanities, <http://dch.phil-fak.uni-koeln.de/index.html>.

15 Cologne Centre for the eHumanities, <http://cceh.uni-koeln.de/>.

16 Simone Kronenwett, *Forschungsdaten an der Philosophischen Fakultät der Universität zu Köln*, Masterarbeit, 2016, <http://nbn-resolving.de/urn:nbn:de:hbz:79pbc-opus-9250>.

17 Sabine Lange-Mauriège 2017, priv comm.

möglich, Projektmittel für FDM-Infrastrukturen direkt bei der Antragstellung zu beantragen und somit das FDM direkt in der Struktur des Projektes zu verankern. Beispielsweise wurden so bereits in den derzeitigen Exzellenz-Einrichtungen der UzK und den im Zukunftskonzept geförderten Centers of Excellence unterstützende FDM-Infrastrukturen etabliert. Des Weiteren werden derzeit im Rahmen von Sonderforschungsbereichen entsprechende Informationsinfrastruktur-Projekte an der UzK gefördert. Ähnliche FDM-Einzel- oder Verbundprojekte wurden im Kontext von BMBF Förderungen zum Aufbau von FDM-Infrastrukturen gefördert.

Beim Aufbau und der Etablierung eines universitätsweiten FDM an der UzK ist die Einbindung der bestehenden, gewachsenen Strukturen notwendig und nützlich. Ebenfalls ist ein Austausch, Abstimmung und die Kooperation mit den bestehenden FDM-Infrastrukturen essentiell. Für eine Zusammenarbeit ist es wichtig, Schnittstellen und Workflows zwischen den zentralen und dezentralen Strukturen zu schaffen. Bestehende Infrastrukturen sollen in ein größeres, universitätsweites Netzwerk integriert werden, um Mehrfacharbeit und Parallelstrukturen zu vermeiden und Synergien zwischen den Strukturen zu schaffen.

3.2. Erarbeitung eines Konzeptes für ein universitätsweites FDM an der UzK

Ausgehend vom Status quo, unter Einbeziehung des Erfahrungsaustausches mit anderen Standorten (u.a. Digitale Hochschule NRW, Georg-August-Universität Göttingen, RWTH Aachen) und unter Berücksichtigung der einschlägigen Empfehlungen von HRK, Allianz-Initiative, RFII, usw. zum FDM (vgl. Abschnitt 1), hat die Projektgruppe ein Konzept für ein universitätsweites FDM entwickelt. Dabei stehen generische, breit einsetzbare Services im Vordergrund. Zurzeit geht es noch vorrangig um Sensibilisierung der Forschenden, der zentralen Einrichtungen, der Universitätsverwaltung, usw. Dabei wurde die Erstellung und Verabschiedung einer universitätsweiten Forschungsdatenleitlinie als ein wichtiges hochschulpolitisches Signal in die Universität hinein angesehen.

3.2.1 Erarbeitung und Verabschiedung einer Leitlinie zum Umgang mit Forschungsdaten an der UzK

In den Empfehlungen der HRK von November 2015 für Hochschulleitungen zur Entwicklung eines FDMs¹⁸ werden Orientierungspfade, Handlungsoptionen, Szenarien und Umsetzungsmöglichkeiten aufgezeigt, die sich beim Auf- oder Ausbau eines institutionellen FDM ergeben. Die HRK bekräftigt, „dass ein wichtiger Erfolgsfaktor darin liegt, dass die Hochschulleitung – nach einer internen Beratung mit allen Interessenvertretern – zu Beginn des Prozesses ein klares, vor allem nach innen gerichtetes Zeichen setzt und damit deutlich macht, dass das FDM künftig ein zentraler Bestandteil des Selbstverständnisses der Hochschule darstellt.“ Die Veröffentlichung einer sogenannten ‘Forschungsdaten-Policy’ im Sinne einer strategischen Leitlinie für das FDM hat sich als bewährtes Mittel erwiesen um das Thema in der Hochschule zu platzieren. Solche Leitlinien können nur dann später Wirksamkeit entfalten, wenn sie zuvor in einem partizipativen Prozess mit den Wissenschaftlerinnen und Wissenschaftlern der Hochschule erarbeitet und abgestimmt werden.

18 „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können.“ HRK.

Im Rahmen der Machbarkeitsstudie für ein universitätsweites FDM an der UzK wurde eine Leitlinie für den Umgang mit Forschungsdaten ausgearbeitet mit Blick auf bestehende Leitlinien deutscher Hochschulen (u.a. Georg-August-Universität Göttingen, Humboldt-Universität zu Berlin, Universität Heidelberg)¹⁹. Diese wurde in einem iterativen Prozess mit den Fakultäten der UzK und anderen Einrichtungen bzw. Interessenvertretern (z.B. Justitiariat) diskutiert und überarbeitet, um eine allgemeine Akzeptanz zu gewährleisten. Diese Leitlinie wurde nach Befürwortung durch das Lenkungsgremium der Machbarkeitsstudie zunächst vom Rektorat und schließlich vom Senat der UzK im Oktober 2017 verabschiedet und im Januar 2018 veröffentlicht.²⁰ Die Leitlinie zum Umgang mit Forschungsdaten an der UzK ist ein wichtiger Baustein für die Etablierung eines standortweiten FDM-Angebots durch die zentralen Infrastruktureinrichtungen.

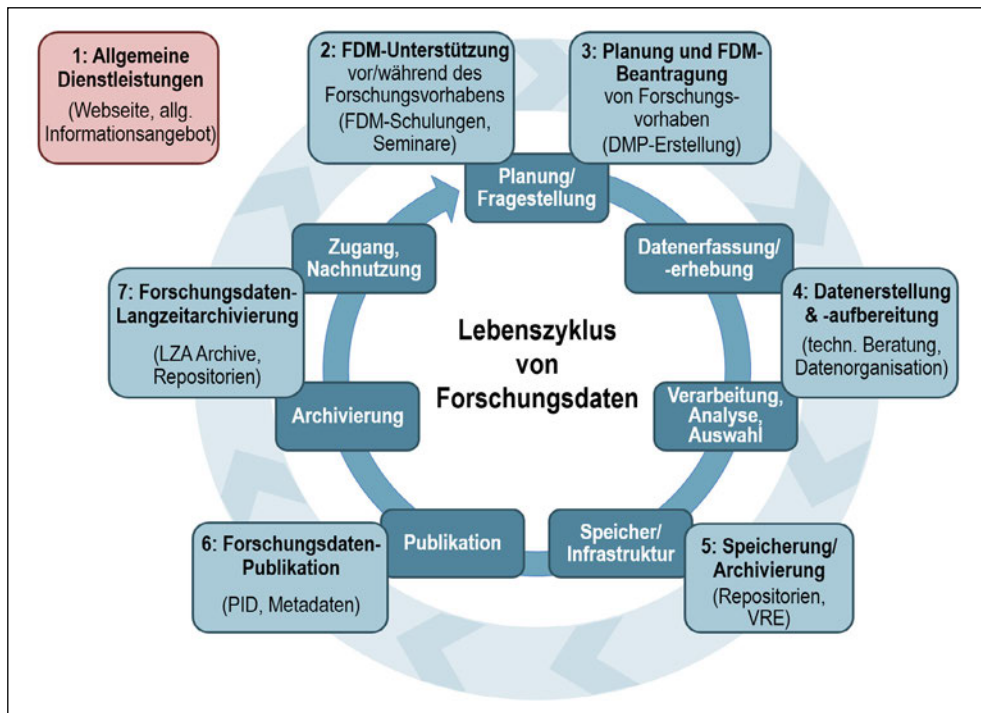


Abb. 1: Mögliche FDM-Dienstleistungen anhand des Lebenszyklus von Forschungsdaten. (Quelle Lebenszyklus: verändert nach Ludwig & Enke²¹, UK Data Archive²²)

19 „Data Policies,“ zuletzt geprüft am 05.03.2018, http://www.forschungsdaten.org/index.php/Data_Policies.

20 „Leitlinie zum Umgang mit Forschungsdaten an der Universität zu Köln,“ zuletzt geprüft am 24.05.2018, https://am.uni-koeln.de/e21463/am_mitteilungen/@6/AM_2018-07_Leitlinie-zum-Umgang-mit-Forschungsdaten.pdf.

21 Jens Ludwig und Harry Enke, Hrsg., *Leitfaden zum Forschungsdaten-Management: Handreichungen aus dem WissGrid-Projekt* (Glückstadt: vwh, 2013).

22 „Research data lifecycle,“ UK Data Archive, zuletzt geprüft am 05.03.2018, <https://www.ukdataservice.ac.uk/manage-data/lifecycle>.

Gemäß den Empfehlungen der HRK geht die Leitlinie zum Umgang mit Forschungsdaten an der UzK auf verschiedene Punkte ein. Neben einer Definition für den Begriff Forschungsdaten wird die gemeinsame Verpflichtung für einen möglichst offenen Umgang mit Forschungsdaten betont. Im Rahmen der Leitlinie verpflichtet sich die UzK, verlässliche und effizient handhabbare Strukturen für das FDM zur Verfügung zu stellen sowie hochschulintern Ansprechpartnerinnen und Ansprechpartner für Beratung und Unterstützung zu benennen. Weiterhin werden Grundsätze zur Veröffentlichung, Dokumentation, Speicherung und Archivierung von Forschungsdaten festgehalten sowie die Berücksichtigung ethischer, datenschutz- und urheberrechtlicher Belange gefordert. Im Gegenzug sind die Wissenschaftlerinnen und Wissenschaftler aufgefordert, sich an die Konventionen eines geregelten FDM zu halten, sowie möglichst Datenmanagementpläne zu erstellen (u.a. bei der Beantragung von Drittmitteln). Schließlich wird in der Leitlinie darauf hingewiesen, dass deren Umsetzung die Situation und Besonderheiten der einzelnen Fächerkulturen zu berücksichtigen hat.

3.2.2 Erarbeitung eines FDM-Service-Portfolios für die UzK

In einem nächsten Schritt wurden mögliche Serviceleistungen der involvierten UzK-Einrichtungen erarbeitet, die den Forschenden der UzK zum FDM angeboten werden könnten. Diese Serviceleistungen wurden noch ohne Bewertung und nicht priorisiert zusammengestellt. Sie basieren im Wesentlichen auf der Bestandsaufnahme des Status quo am Standort Köln, dem Erfahrungsaustausch mit anderen Standorten/Netzwerken und einer Literaturrecherche. Hierbei soll kein Anspruch auf Allgemeingültigkeit und Vollständigkeit erhoben werden. Allen Beteiligten war außerdem bewusst, dass es im Rahmen einer ersten Aufbauphase nicht möglich sein würde, ein umfassendes und vollständiges Serviceportfolio zum FDM aufzubauen, sondern dass es notwendig sein würde, sich im Aufbau der Dienstleistungen bewusst auf einzelne Maßnahmen zu konzentrieren, mit denen begonnen werden kann. Insofern war diese Vorarbeit sehr hilfreich, um für den geplanten Aufbau sinnvolle Maßnahmen definieren und priorisieren zu können, die dann als Arbeitspakete angegangen werden können.

Das erarbeitete Service-Portfolio enthält potentielle FDM-Dienstleistungen zur Unterstützung der Wissenschaftlerinnen und Wissenschaftler entlang des Lebenszyklus von Forschungsdaten (Abb. 1). Diese wurden in sieben Dienstleistungsgruppen eingeteilt und sind in Tabelle 2 mit konkreten Inhalten vorgestellt.

Tabelle 2: Umfassendes FDM-Service-Portfolio

1. Übergreifende Dienstleistungen
<ul style="list-style-type: none">• Allgemeine Informationen zum FDM auf eigener FDM-UzK-Webseite präsentieren• E-Learning-Materialien (z.B. Webinare) bereitstellen• Andere Formate zur allgemeinen Information zum FDM aufbauen und etablieren: z. B. FDM
2. Dienstleistungen im Vorfeld eines Forschungsvorhabens (insb. für die Zielgruppe des wiss. Nachwuchses)
<ul style="list-style-type: none">• Aufbau und Etablierung von Schulungen (u.a. Grundlagen- und Vertiefungsschulungen) für verschiedene Zielgruppen wie z.B. Professorinnen und Professoren, Postdocs, Promovierende, Master-Studierende, Verwaltung• „Globale Themen“ für Workshop-/Seminar-Formate wie z.B.: Was ist FDM? Vorteile von FDM? Finanzierung von FDM? Erstellung von Datenmanagementplänen• Unterstützung bei der Entwicklung und Implementierung von Seminaren im Curriculum mit Credit Points
3. Dienstleistungen bei der Planung und gegebenenfalls bei Beantragung eines Forschungsvorhabens
<ul style="list-style-type: none">• Beratungsdienstleistungen zur Antragstellung von Drittmittelanträgen:<ul style="list-style-type: none">– Förderlinien, Fördermöglichkeiten für FDM-Aktivitäten– Unterstützung bei der Budgetierung von FDM– Vorgaben durch die Universität– Verantwortlichkeiten bei FDM• Beratungsleistungen zur Erstellung von DMPs<ul style="list-style-type: none">– Bereitstellung von Templates (je nach Förderlinie) unter Berücksichtigung vorher genannter Beratungspunkte– Beratung zu den Phasen des Forschungsdatenzyklus und zum Inhalt von DMPs– Unterstützung bei der Verwendung von DMP-Tools– Sammlung und Bereitstellung von Best-Practices• Bei Bedarf Bereitstellung von lokal gehosteten DMP Tools (z.B. DMPonline, DMPTool, RDMO)• Unterstützung bei der Erstellung einer Datenstrategie und eines FDM-Finanzierungsplans einzelner Großprojekte<ul style="list-style-type: none">– Beratung zum technischen Projektmanagement für die projektbezogene, individuelle Speicherung von Forschungsdaten, Qualitätssicherung, Prozessabläufe, Data Life Cycle, Versionierung, Rechtemanagement & Zugriffskontrolle, Metadaten und Metadaten-service der Uni– Beratung zur Kooperation mit anderen Einrichtungen (intern/extern/Verbünde), Data Sharing, Data Conversion, Formate

4. Dienstleistungen bei der Datenerstellung und der Datenaufbereitung

- In der aktiven Forschungsphase laufen alle Services aus den anderen Bereichen zusammen:
- Beratung zu Dateiformaten, Datenorganisation, Datenspeicherung
- Mitwirkung bei der Erstellung individualisierter Metadaten-Schema
- Bereitstellung einer kooperativen Arbeitsplattform oder Elementen davon
- Beratung zu Datenaufbereitung, Dokumentation, Metadaten

5. Dienstleistungen bei der Publikation von Forschungsdaten

- Beratung zu Datenschutz, Urheber- und Nutzungsrechten, Nachnutzung, Lizenzierung (z.B. Creative-Commons)
 - Beratung zu und Vergabe von Persistenten Identifikatoren (PIDs, z.B. DOI), zum Publizieren und Zitieren von Daten
 - Beratung zum Ingest-Verfahren bei externen (fachlichen überregionalen) Daten-Repositories, gegebenenfalls auch Hilfestellung beim Daten-Ingest
 - Verknüpfung der Metadaten mit den zugehörigen Publikationsdaten
 - Bereitstellen von Schnittstellen (OAI) zum Harvesten von Metadaten in überregionalen Datenbanken
 - Erschließung der Metadaten in einer Nachweisdatenbank / Hochschulbibliographie, Verknüpfung mit Forschungsinformationssystem
 - Kontrolle und Korrektur von Metadaten
 - Verknüpfung der Metadaten mit den Normdaten der Deutsche Nationalbibliothek und vorhandenen Autoren-Identifikatoren (z.B. ORCID)
 - Bereitstellen eines Rechtemanagements zur Regelung des Daten-Zugriffs und eventueller Embargozeiten
 - Organisation der Pflichtabgabe an die Deutsche Nationalbibliothek (sofern erforderlich)
 - Integration der Daten in ein Forschungsdatenrepositorium
 - Bereitstellung einer individualisierten Präsentationsoberfläche
-

6. Dienstleistungen zur Speicherung und Archivierung der Daten

Storage für alle gewünschten Domänen (privat, Gruppe, Dauer, öffentlich)²³ gemäß HRK-Empfehlungen.²⁴ Folgende Szenarien sind zu erwarten:

- Private Domäne: Einfache Abschlussarbeiten, Einzelprojekte (DFG, Uni, usw.)
 - Zentrales universitätsweites Repositorium für kleine Datenmengen oder Abschlussarbeiten (Long-Tail-Archiv), mit PIDs, Kosten überschaubar, Aufbewahrungszeit 10 Jahre (ggf. zunächst unbeschränkt), zentral organisiert
- Gruppendomäne: Projekte in stark vernetzter Community
 - Beratung bei der Nutzung externer (intern-)nationaler Angebote, externe Vorgaben
 - Forschungsdaten über das Projekt verteilt, Zugriffsverwaltung
 - Storage aus Institutsmitteln oder von Universität zentral gestellt, Backup & Langzeitarchivierung über Universität geregelt
- Dauerhafte Domäne & öffentlich: Universitäre Verbünde & internationale Kollaborationen (SFB, Exzellenz-Cluster, EU-Projekte):
 - Beratung bei Aufbau eines lokalen Repositoriums
 - (Heterogene) Forschungsdaten werden lokal vorgehalten, sind von außen erreichbar
 - Bereitstellung von zentralem Storage, Backup über die Universität geregelt
- Industriekooperation:
 - Beratung bei Sicherheitsfragen, Zugriffsschutz
 - Bereitstellung von Storage, Backup & Archivierung auf Wunsch

Eine generische Service-Architektur soll gemäß HRK-Empfehlungen „zentrale Formen der Datenerhaltung“ unterstützen:

- Technische Plattform (Hardware, Software, Access), Werkzeuge, Dienste für die Bereitstellung von Storage, Cloud-Speicher
- Technische Plattform für (Langzeit-)archivierung
- Virtualisierungsplattform für Nutzung von Storage, Administration durch Universität als fertige Lösung für Endkunden, ggf. auf Cloud-Basis
- Zentrale FDM-Infrastruktur für kleine Projekte (s.o.) mit Funktionen wie bspw. Daten-Ingest, Metadatenbeschreibung, Präsentation von Daten, Backup und Archivierung, PIDs, Software Angebote (z.B. Microsoft SharePoint) usw.

7. Dienstleistungen zur Langzeitarchivierung der Forschungsdaten

- Formatvalidierung als Basis für die Risikoabschätzung und ggf. Maßnahmenplanung für die Langzeitverfügbarkeit
- Übernahme der Daten in ein Langzeitarchivierungs-System (z.B. Rosetta)

23 Andrew Treloar, David Groenewegen und Cathrine Harboe-Ree, „The Data Curation Continuum: Managing Data Objects in Institutional Repositories“, *D-Lib Magazine* 13, Nr. 9/10 (September 2007), <https://doi.org/10.1045/september2007-treloar>.

24 „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können.“ HRK.

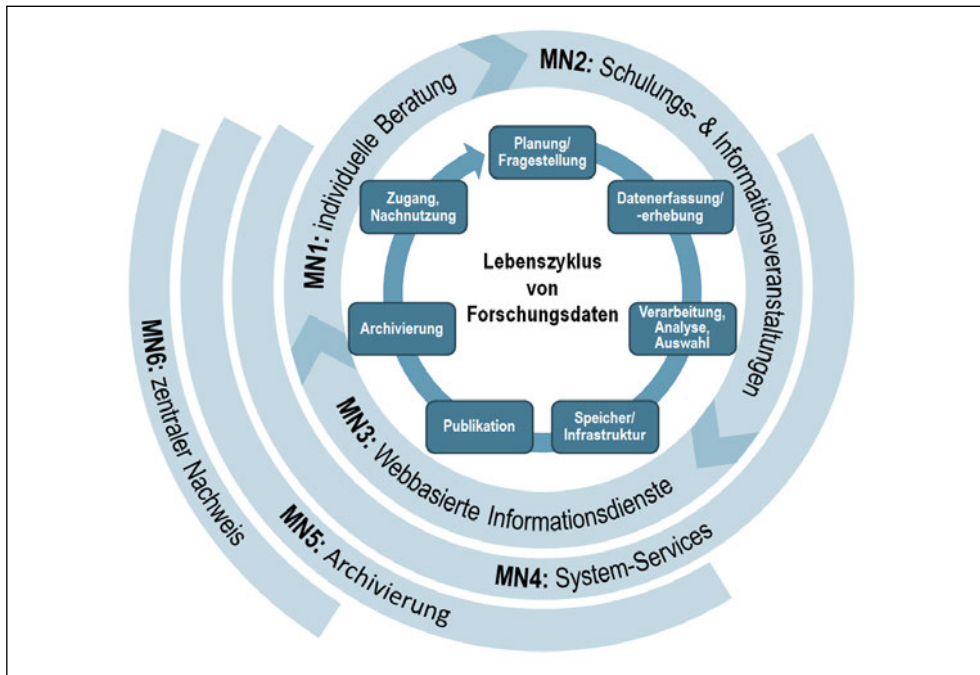


Abb. 2: Abdeckung des Lebenszyklus von Forschungsdaten durch die geplanten Maßnahmenpakete. (Quelle Lebenszyklus: verändert nach Ludwig & Enke²⁵, UK Data Archive²⁶)

3.2.3 Erstellung eines Maßnahmenkatalogs für den Aufbau eines FDM-Angebots

In einem nächsten Arbeitsschritt wurde anhand des erarbeiteten, umfangreichen FDM-Service-Portfolios, das ein langfristiges Entwicklungsziel darstellt, ein Katalog an Maßnahmen abgeleitet. Für den mittelfristigen Aufbau und die Etablierung einer FDM-Unterstützungsinfrastruktur ist in einem Entwicklungsszenario eine Priorisierung von Basisdienstleistungen bis hin zu anspruchsvolleren, spezifischeren Diensten zu treffen. Anhand der genannten Bedürfnisse der Forschenden und den daraus abgeleiteten Anforderungen wurde ein entsprechender Maßnahmenkatalog aus dem Service-Portfolio erarbeitet, der in zukünftigen Projekten umgesetzt werden könnte. Die Projektpartner haben sich zunächst den Aufbau von Basisdienstleistungen zum Ziel gesetzt, die eine möglichst breite Nutzerschaft ansprechen. In weiteren Projekten könnten diese Basisservices zu einem umfassenderen Serviceangebot ausgebaut werden.

Neben den inhaltlichen Aspekten ist die Koordinierung des Serviceangebotes und die Einbettung in bestehende Organisationsstrukturen der Uzk von zentraler Bedeutung. Mit dem Ziel der Bündelung von Ressourcen und Kompetenzen zum FDM soll ein Service-Zentrum eingerichtet werden, welches

25 Ludwig und Enke, *Leitfaden zum Forschungsdaten-Management*.

26 „Research data lifecycle.“ UK Data Archive, zuletzt geprüft am 05.03.2018, <https://www.ukdataservice.ac.uk/manage-data/lifecycle>.

den Forschenden an der UZK als zentraler Ansprechpartner für FDM-Themen zur Verfügung steht. Das Zentrum wird zunächst durch die drei beteiligten Einrichtungen Dezernat Forschungsmanagement (D7), Regionales Rechenzentrum (RRZK) und Universitäts- und Staatsbibliothek (USB) gebildet. Die Umsetzung der Maßnahmenpakete erfolgt durch die drei Projektpartner.

Ausgehend von der vorhandenen Expertise der Projektpartner und infrastrukturellen Lösungen an der Universität bilden auf der einen Seite Entwicklungen in den Bereichen Beratung, Schulung und Information (Sensibilisierung mit dem Thema FDM) inhaltliche Schwerpunkte der Maßnahmenpakete (MN1-3). Neben USB und RRZK ist das D7 schwerpunktmäßig in diese Maßnahmenpaketen involviert. Ein zweiter Schwerpunktbereich sind digitale Dienste/Werkzeuge, die die Forschenden in ihren Arbeitsprozessen einsetzen können (MN4-6). Diese werden im Wesentlichen von RRZK und USB behandelt. Zurzeit ist unsere Erfahrung, dass sich Forschenden aus den Geistes- und Gesellschaftswissenschaften eher an die Bibliothek und Forschende aus den Naturwissenschaften eher an das Rechenzentrum wenden. Wir erwarten allerdings, dass sich dieses Verhalten mit dem Aufbau des Service-Zentrums als zentrale Anlaufstelle ändern und somit eine effizientere Arbeitsverteilung möglich sein wird. Abb. 2 zeigt, wie die einzelnen Maßnahmenpakete die verschiedenen Phasen des Forschungsdaten-Lebenszyklus bzw. des Forschungsprozesses abdecken. Über den gesamten Zyklus hinweg werden zentrale, skalierbar einsetzbare Services bereitgestellt. Es wurden sechs Maßnahmenpakete mit spezifischen Zielen identifiziert (Tabelle 3). Bei einer möglichen Übertragung der Maßnahmen auf andere Standorte ist natürlich die Aufstellung der Informationsinfrastrukturdienstleister und insbesondere die Kompetenzverteilung zu berücksichtigen.

Tabelle 3: Übersicht der Maßnahmen mit spezifischen Zielen und Aufgaben

Maßnahme 1 (MN1): Individuelle Beratung	
Ziel:	Die Forschenden an der UzK erhalten für alle Fragen des FDM entlang des gesamten Forschungszyklus individuelle Beratung und Hilfestellung
Aufgaben:	Individuelle Antragsberatung bei Drittmittelprojekte (u.a. Überprüfung von Anträgen nach FDM-Kriterien, Erstellung von Datenmanagementplänen und FDM-Antragstexten) Unterstützung bei Begutachtungen von Drittmittelprojekten Austausch mit dezentralen FDM-Beratungsstellen Awareness schaffen und Austausch mit Abteilungen innerhalb des RRZK, USB und D7 zu aktuellen FDM-Themen, Fördergeberrichtlinien, etc. Beratung zum technisches Datenmanagement (u.a. Anlaufstelle für technische Erstberatung)
Maßnahme 2 (MN2): Schulungs- und Informationsveranstaltungen	
Ziel:	Es wird ein umfassendes Programm an Schulungs- und Informationsveranstaltungen (insbes. für Nachwuchswissenschaftlerinnen und -wissenschaftler, Promovenden, Studierende vor Abschlussarbeiten, aber auch für alle anderen Forschenden) an der UzK angeboten
Aufgaben:	Planung, Organisation und Durchführung von zielgruppenspezifischen FDM-Informationsveranstaltungen, Grundlagenschulungen und Vertiefungsschulungen Erstellung von Schulungsmaterialien
Maßnahme 3 (MN3): Webbasierte Informationsdienstleistungen zum Forschungsdatenmanagement	
Ziel:	Interessierte Forschende finden auf der Webseite des Service-Zentrums vielfältige, modular aufgebaute Informationsangebote zu allen Aspekten des FDM, die sie zielgerichtet und just-in-time zu ihren spezifischen Fragestellungen nutzen können.
Aufgaben:	Aufbau und Pflege eines webbasierten Informationssystems Erarbeitung von FDM-Inhalten (u.a. Tutorien, Webinare) Bereitstellung von DMP-Tools wie RDMO
Maßnahme 4 (MN4): System-Dienste für Nutzer des Forschungsdatenmanagements	
Ziel:	Einzelne Wissenschaftler und Wissenschaftlerinnen sowie verschiedene Forschungsverbünde und Forschungsprojekte können Informationsinfrastruktur-Angebote nutzen.
Aufgaben:	Mitwirkung bei der Erstellung individualisierter Metadatenschemata Projektbegleitung zu speziellen FDM-Themen Bereitstellung und Verwaltung von Speicher für die gewünschten Domänen (Privat, Gruppe, Dauerhaft, Öffentlich) für Datenspeicherung und Verarbeitung Bereitstellung und Betreuung von kooperativen Arbeitsplattformen für Verbundprojekte

Maßnahme 5 (MN5): Archivierung von Forschungsdaten

Ziel: Die Forschenden erhalten Informationen, Beratung und Entscheidungshilfen zur lokalen und externen Archivierung ihrer Forschungsdaten. Insbesondere für Long-Tail- Forschungsprojekte steht eine institutionelle Archivierungs-Infrastruktur zur Verfügung.

Aufgaben: Backup und befristete Archivierung von Daten
Bereitstellung eines PID-Service (z.B. DOI, Handle)
Maßnahmen für die Langzeitarchivierung (z.B. TSM oder Rosetta)

Maßnahme 6 (MN6): Aufbau einer zentralen Nachweisdatenbank für (veröffentl.) Forschungsdaten

Ziel: Alle an der UzK veröffentlichten Forschungsdaten sind in einer zentralen Nachweisdatenbank erfasst und somit international recherchierbar. Diese Daten finden sich auch in der geplanten Hochschulbibliographie sowie dem geplanten Forschungsinformationssystem, und stehen somit auch für das Berichtswesen der UzK zur Verfügung.

Aufgaben: Auswahl und Bewertung von Metadaten Schemen
Auswahl und Implementation einer technischen Plattform für die zentrale Nachweisdatenbank incl. Entwicklung einer Eingabemaske für die Forschenden
Laufende Erschließung und Bearbeitung von Selbstmeldungen

4. Zusammenfassung und Ausblick

Die Machbarkeitsstudie hat gezeigt, dass es an der UzK bereits einzelne Initiativen zu einem spezifischen FDM an Fakultäten, an Instituten und innerhalb von Forschungsverbänden gibt. Diese agieren weitestgehend eigenständig und unabhängig von den anderen Informationsinfrastrukturen. An der UzK erfolgt noch kein standortweites, systematisches FDM. Ein solches ist jedoch, wie dargelegt, für die Antragsfähigkeit für Drittmittel (insb. Exzellenz-Strategie, Sonderforschungsbereiche), aber auch für den Long-Tail,²⁷ von immer größerer Bedeutung. Die Studie hat dabei zwei Handlungsfelder für einen Auf- und weiteren Ausbau der Informationsinfrastruktur aufgezeigt: (i) Beratung, Information und Schulung und (ii) technische FDM-Infrastruktur. Dabei geht es im Wesentlichen sowohl um eine stärkere Vernetzung der laufenden Unternehmungen, als auch um die Identifizierung und Entwicklung von skalierbaren zentralen Basis-FDM-Dienstleistungen für die gesamte Universität. Ein verbesserter Erfahrungs- und Wissensaustausch schafft zudem die Möglichkeit, Synergien auf den verschiedensten Ebenen bei der Weiterentwicklung der Standort-Infrastruktur zu nutzen.

Als Ergebnis wurden im Bereich der zwei Handlungsfelder sechs Maßnahmenpakete für die Etablierung einer Unterstützung für ein universitätsweites systematisches FDM entwickelt. Ziel ist es hierbei, eine zentrale Kontaktstelle für Fragen zum FDM in allen Phasen eines Forschungsvorhabens für die Universität zu bieten. Ganz essentiell wird dabei die Einbeziehung der existierenden spezifischen Strukturen in das zu bildende Expertennetzwerk sein, welches um allgemeine zentrale Beratungs-

²⁷ P. Bryan Heidorn, „Shedding Light on the Dark Data in the Long Tail of Science,“ *Library Trends* 57, Nr. 2 (27. März 2009): 280–299, <https://doi.org/10.1353/lib.0.0036>. Christine L. Borgman u.a., „Data Management in the Long Tail: Science, Software, and Service,“ *International Journal of Digital Curation* 11, Nr. 1 (12. Oktober 2016): 128–148, <http://doi.org/10.2218/ijdc.v11i1.428>.

Schulungs- und digitale Dienstleistungen ergänzt werden soll. Dafür bringen die drei zentralen Einrichtungen das Dezernat Forschungsmanagement (D7), das Regionale Rechenzentrum (RRZK) und die Universitäts- und Stadtbibliothek (USB) der Universität zu Köln ihre Expertise auf den Gebieten Wissenschaftsmanagement und Wissensinfrastruktur in ein „Service-Zentrum Forschungsdaten“ ein. Dieses soll mit seinen Services in den folgenden Jahren aufgebaut und etabliert werden. Unter anderem soll dabei auch ein Augenmerk darauf gelegt werden, welche Personenkreise (z. B. höherer oder gehobener Dienst, Fachreferentinnen und -referenten in der USB, Fakultätsbetreuerinnen und -betreuer im RRZK) innerhalb der Projektpartner mit in die Strukturen aufgenommen werden können.

Der Bedarf, solche Unterstützungsstrukturen aufzubauen, wird durch mehrere Anfragen von Forschenden zu Publikation, Archivierung und Berücksichtigung in Drittmittelanträgen bereits während der Laufzeit der Machbarkeitsstudie bestätigt.

In einem nächsten Schritt muss für die Umsetzung eine geeignete Organisations- bzw. Kooperationsstruktur gefunden werden. Dabei wird die hohe Autonomie der Fakultäten eine wichtige Rolle spielen. Denkbar wäre ein Netzwerk, das u.a. eine Differenzierung und Spezialisierung in Knotenpunkten innerhalb der Fachbereiche ermöglicht. Zudem haben sich moderierte Workshops als geeignete Formate in Konfliktsituationen erwiesen. In enger Zusammenarbeit mit den institutionellen Einheiten (z.B. Data Center for the Humanities), Forschungsverbänden und der Universitätsleitung sollen Konzepte für ein umfassenderes, vernetztes FDM an der Universität, welche sowohl Fach- und Projektspezifika als auch Kostenmodelle berücksichtigen, ausgearbeitet und bereits vereinzelt exploriert werden.

Danksagung

Die Autoren danken den Kolleginnen und Kollegen der Projektpartner Claudia Berchtold (D7), Johannes Boll (RRZK), Ralf Depping (USB), Sonja Kloppenburg (D7), Maria Riese (D7), Matthias Röder (D7) und Volker Winkelmann (RRZK) für die Durchführung der Machbarkeitsstudie, die gemeinsame Erarbeitung der Ergebnisse und für konstruktives Feedback bei der Erstellung dieses Beitrags. Die Autoren danken den wertvollen Kommentaren der Gutachterinnen und Gutachter, die zur Qualitätssteigerung dieses Artikels beigetragen haben.

Literaturverzeichnis

- Borgman, Christine L., Milena S. Golshan, Ashley E. Sands, Jillian C. Wallis, Rebekah L. Cummings, Peter T. Darch und Bernadette M. Randles. „Data Management in the Long Tail: Science, Software, and Service.“ *International Journal of Digital Curation* 11, Nr. 1 (12. Oktober 2016): 128–148. <https://doi.org/10.2218/ijdc.v11i1.428>.
- Dierkes, Jens und Ulrike Wuttke. „The Göttingen eResearch Alliance: A Case Study of Developing and Establishing Institutional Support for Research Data Management.“ *ISPRS International Journal of Geo-Information* 5, Nr. 8 (1. August 2016): 133. <https://doi.org/10.3390/ijgi5080133>.

- Eifert, Thomas, Stephan Muckel und Dominik Schmitz. „Introducing research data management as a service suite at RWTH Aachen university.“ In 9. *DFN-Forum Kommunikationstechnologien*, P-257:55–64. Lecture Notes in Informatics. Gesellschaft für Informatik, Bonn, 2016. <http://subs.emis.de/LNI/Proceedings/Proceedings257/article10.html>.
- Flanders, David F. „A New Model for Research IT Services.“ *David F. Flanders* (Blog), 15. April 2013. <https://dfflanders.wordpress.com/2013/04/15/resbaz/>.
- Gertz, Michael, Carlos Rueda und Jianting Zhang. „Interoperability and Data Integration in the Geosciences.“ In *Scientific Data Management: Challenges, Technology, and Deployment*, herausgegeben von Arie Shoshani und Doron Rotem, 369–398. Computational Science Series. CRC Press/Chapman & Hall, 2009.
- Heidorn, P. Bryan. „Shedding Light on the Dark Data in the Long Tail of Science.“ *Library Trends* 57, Nr. 2 (27. März 2009): 280–299. <https://doi.org/10.1353/lib.0.0036>.
- HRK, Hochschulrektorenkonferenz. „Management von Forschungsdaten–eine zentrale strategische Herausforderung für Hochschulleitungen. Empfehlung der 16. Mitgliederversammlung der HRK am 13. Mai 2014 in Frankfurt am Main.“ Zuletzt geprüft am 23.05.2018. <https://www.hrk.de/positionen/beschluss/detail/management-von-forschungsdaten-eine-zentrale-strategische-herausforderung-fuer-hochschulleitungen/>.
- HRK, Hochschulrektorenkonferenz. „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können: Orientierungspfade, Handlungsoptionen, Szenarien. Empfehlung der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel.“ Zuletzt geprüft am 23.05.2018. <https://www.hrk.de/positionen/beschluss/detail/wie-hochschulleitungen-die-entwicklung-des-forschungsdatenmanagements-steuern-koennen-orientierungsp/>.
- Johnston, Lisa, Jake Carlson, Patricia Hswe, Cynthia Hudson-Vitale, Heidi Imker, Wendy Kozlowski, Robert Olendorf und Claire Stewart. „Data Curation Network: How Do We Compare? A Snapshot of Six Academic Library Institutions’ Data Repository and Curation Services.“ *Journal of eScience Librarianship* 6, Nr. 1 (28. Februar 2017). <https://doi.org/10.7191/jeslib.2017.1102>.
- Kindling, Maxi, Peter Schirmbacher und Elena Simukovic. „Forschungsdatenmanagement an Hochschulen: Das Beispiel der Humboldt-Universität zu Berlin“. *LIBREAS. Library Ideas*, Nr. 23 (2013). <http://doi.org/10.18452/9041>.
- Kormeier, Benjamin, Klaus Hippe und Ralf Hofestädt. „Data Warehouses in Bioinformatics: Integration of Molecular Biological Data.“ *it - Information Technology Methoden und innovative Anwendungen der Informatik und Informationstechnik* 53, Nr. 5 (2011): 241–249. <https://doi.org/10.1524/itit.2011.0649>.

- Kratzke, Jonas und Vincent Heuveline. *E-Science-Tage 2017: Forschungsdaten managen*. Heidelberg: heiBOOKS, 2017. <https://doi.org/10.11588/heibooks.285.377>.
- Kronenwett, Simone. *Forschungsdaten an der Philosophischen Fakultät der Universität zu Köln*. Masterarbeit, 2016. <http://nbn-resolving.de/urn:nbn:de:hbz:79pbc-opus-9250>.
- Ludwig, Jens und Harry Enke, Hrsg. *Leitfaden zum Forschungsdaten-Management: Handreichungen aus dem WissGrid-Projekt*. Glückstadt: vwh, 2013.
- Pryor, Graham, Sarah Jones und Angus Whyte. *Delivering research data management services: Fundamentals of good practice*. Facet Publishing, 2013.
- Rat für Informationsinfrastrukturen. *Leistung aus Vielfalt: Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland*. Göttingen, 2016. <http://nbn-resolving.de/urn:nbn:de:101:1-201606229098>.
- Ray, Joyce M. *Research data management: Practical strategies for information professionals*. Purdue University Press, 2014.
- Treloar, Andrew, David Groenewegen und Cathrine Harboe-Ree. „The Data Curation Continuum: Managing Data Objects in Institutional Repositories.“ *D-Lib Magazine* 13, Nr. 9/10 (September 2007). <https://doi.org/10.1045/september2007-treloar>.

Datenmanagementpläne leicht gemacht: TUB-DMP, das Web-Tool der TU Berlin für Datenmanagementpläne

Monika Kuberek und Fabian Fürste, Universitätsbibliothek der Technischen Universität Berlin

Zusammenfassung:

Das Web-Tool TUB-DMP ist ein weiterer Baustein in der Forschungsdaten-Infrastruktur der Technischen Universität Berlin, die vom „Servicezentrum Forschungsdaten und -publikationen“ (SZF) betrieben wird. TUB-DMP ergänzt das Repositorium für Forschungsdaten und Publikationen „Deposit-Once“, das die langfristige Verfügbarkeit, Verifizierung, Zitierbarkeit und Nachnutzbarkeit der Forschungsergebnisse gewährleistet. Forschungsdatenmanagement und Forschungsdaten-Infrastruktur der TU Berlin sind nach den FAIR Data Principles (Findable, Accessible, Interoperable, Re-usable) ausgerichtet. Dieser Maßgabe entsprechend wurde auch TUB-DMP entwickelt, um den Wissenschaftlerinnen und Wissenschaftlern der TU Berlin ein Dokumentationswerkzeug an die Hand zu geben, das ein FAIRes Management ihrer Forschungsdaten ermöglicht. TUB-DMP wurde gemeinsam mit Pilotpartnern aus der TU Berlin entwickelt. Die erste Version wurde im September 2015 in Betrieb genommen, Version 2 mit einer Reihe von Erweiterungen im Dezember 2017.

Summary:

The TUB-DMP Web tool is a further component in the research data infrastructure of the Technische Universität Berlin (TU Berlin), which is operated by the Service Center for Research Data and Publications (SZF). TUB-DMP complements “DepositOnce”, the repository for research data and publications, which guarantees the long-term availability, verification, citability and re-usability of research results. The management of research data and the research data infrastructure at TU Berlin are aligned with the FAIR data principles (Findable, Accessible, Interoperable, Re-usable). TUB-DMP was developed according to these requirements in order to provide scientists with a documentation tool that enables the FAIR data management of their research data. TUB-DMP was developed in cooperation with pilot partners from TU Berlin. The first version went live in September 2015, the extended second version was released in December 2017.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S47-60>

Autorenidentifikation: Kuberek, Monika: ORCID: <https://orcid.org/0000-0002-1672-5271>; Fürste, Fabian: ORCID: <https://orcid.org/0000-0003-2172-609X>

Schlagwörter: Datenmanagementplan, DMP, FAIR Data Principles, Forschungsdatenmanagement, FDM

1. TUB-DMP – ein Web-Tool für Datenmanagementpläne

Das Web-Tool TUB-DMP ist ein weiterer Baustein in der Forschungsdaten-Infrastruktur der Technischen Universität Berlin, die vom Servicezentrum Forschungsdaten und -publikationen (SZF)¹ betrieben wird. TUB-DMP ergänzt das Repositorium für Forschungsdaten und Publikationen DepositOnce,² das die langfristige Verfügbarkeit, Verifizierung, Zitierbarkeit und Nachnutzbarkeit der Forschungsergebnisse (= Forschungsdaten und Publikationen) gewährleistet. Forschungsdatenmanagement (FDM) und Forschungsdaten-Infrastruktur der TU Berlin sind nach den FAIR Data Principles³ (Findable, Accessible, Interoperable, Re-usable) ausgerichtet. Dieser Maßgabe entsprechend wurde auch TUB-DMP entwickelt, um den Forschenden an der TU Berlin ein Dokumentationswerkzeug an die Hand zu geben, das ein FAIRes Management ihrer Forschungsdaten ermöglicht.

1.1. Was ist ein Datenmanagementplan (DMP)?

Ein DMP trägt dazu bei, dass Forschungsdaten langfristig ihren Wert behalten. Er dokumentiert den Kontext, in dem die Forschungsdaten entstanden sind, sowie den Umgang mit den Forschungsdaten während ihrer Entstehung und nach Beendigung ihrer Erhebung. Der DMP enthält Angaben zu Projektleitung, Projekttitel, Projektbeschreibung, Förderorganisation, Art und Umfang der entstandenen Forschungsdaten, Speicherung, Aufbewahrungsdauer usw. Diese Informationen dienen insbesondere dazu, Forschungsdaten in späteren Jahren interpretieren, nachvollziehen und somit nachnutzen zu können.

Die Verfügbarkeit gut organisierter und gepflegter Daten gehört heute zur guten wissenschaftlichen Praxis und gewährleistet die Transparenz der Forschungsergebnisse, eine grundlegende Anforderung im Forschungsbereich.

1.2. Ausgangslage

Im letzten Jahrzehnt ist weltweit die Erkenntnis gewachsen, dass Forschungsdaten eine unschätzbare Ressource und einen Wert an sich darstellen, den es zu erhalten, disziplinübergreifend bereitzustellen und nachzunutzen gilt. Das in der TU Berlin in den Jahren 2012–2014 aufgebaute und seitdem etablierte FDM umfasst den gesamten Lebenszyklus der Forschungsdaten und steht mit seiner konzeptionellen Zielstellung – Speicherung, Verifizierung, Langzeitarchivierung und Nachnutzung von Forschungsergebnissen – auch dafür ein, dass die Forschungsdaten langfristig ihren Wert als Forschungsressource behalten.⁴ So sollen in DepositOnce, dem Repositorium der TU Berlin, zusammen mit den Forschungsdaten auch Informationen gespeichert werden, die zur Verifizierung und Reproduktion eines Forschungsergebnisses notwendig sind.⁵

1 Das SZF ist eine gemeinsame Einrichtung der Universitätsbibliothek (UB), des IT-Service-Center tubIT und der Forschungsabteilung (Abt. V); die Leitung des SZF ist in der Hauptabteilung 3 „Elektronische Dienste“ (HA 3) der UB angesiedelt, zuletzt geprüft am 15.02.2018, http://www.szf.tu-berlin.de/menue/ueber_das_szf.

2 Repositorium DepositOnce, zuletzt geprüft am 15.02.2018, <https://depositonce.tu-berlin.de>.

3 Mark D. Wilkinson et al., „The FAIR Guiding Principles for scientific data management and stewardship,“ *Sci. Data* 3:160018 (2016), <https://doi.org/10.1038/sdata.2016.18>.

4 Monika Kuberek, *Organisatorisch-technisches Konzept für eine Forschungsdaten-Infrastruktur in der TU Berlin*. Technische Universität Berlin, 15. März 2012, 4–5, <http://dx.doi.org/10.14279/depositonce-6604>.

5 Ebd., 14–15.

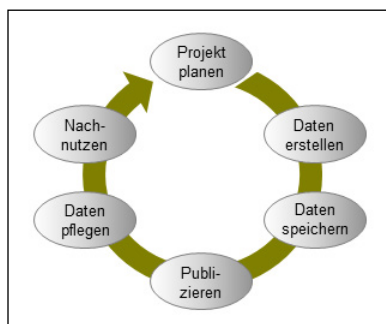


Abb. 1: Forschungsdaten-Lebenszyklus (Grafik angelehnt an: <https://www.ukdataservice.ac.uk/manage-data/lifecycle>)

Darüber hinaus sind für die Nachnutzung der Forschungsdaten weitere Informationen unerlässlich, wie zum Beispiel: Wie und in welchem Rahmen / Projekt sind die Forschungsdaten entstanden? Welche Methoden wurden bei der Erstellung angewendet? Welches ist der Erhebungszeitraum? Wie und wo werden die Daten gespeichert? Welche Kosten entstehen langfristig? Der Bereitstellung dieser Informationen dient der DMP.

Vor dem Hintergrund der Nachnutzung von Forschungsdaten gehört die Erstellung von DMPs in den anglo-amerikanischen Ländern seit etlichen Jahren zum Standard bei der Projektantragstellung; sie werden in aller Regel von den dortigen Förderorganisationen eingefordert. Ebenso wie beim FDM insgesamt kommt Großbritannien auch bei der Entwicklung und dem Einsatz von DMPs die Vorreiterrolle in Europa und eine führende Rolle weltweit zu. Der erste standardisierte DMP⁶ in Großbritannien wurde 2009 vom Digital Curation Centre (DCC)⁷ entwickelt und seitdem kontinuierlich weiterentwickelt. Die DCC-Checklist, die inzwischen in der 4. Version⁸ vorliegt, ist die in der Wissenschaftslandschaft international am weitesten verbreitete Vorlage zur Erstellung eines DMPs; sie war auch das Vorbild für die im DFG-Projekt WissGrid⁹ entwickelte „Checkliste zum Forschungsdaten-Management“.¹⁰

- 6 Martin Donnelly und Sarah Jones, *Data Management Plan Content Checklist: Draft Template for Consultation* (Edinburgh: Digital Curation Centre, 17.06.2009), zuletzt geprüft am 15.02.2018, http://www.dcc.ac.uk/webfm_send/1280. Siehe auch: *Checklist for a Data Management Plan*, zuletzt geprüft am 15.02.2018, <http://www.dcc.ac.uk/resources/data-management-plans/checklist>.
- 7 DCC wurde 2004 auf Empfehlung von JISC (<https://www.jisc.ac.uk/>), dem damaligen „Joint Information Systems Committee“ als nationales Zentrum zur Sicherung des dauerhaften Zugangs und der Langzeitarchivierung digitaler Ressourcen gegründet, zuletzt geprüft am 15.02.2018, <http://www.dcc.ac.uk/about-us/history-dcc/history-dcc>.
- 8 DCC, *Checklist for a Data Management Plan*, v4.0 (Edinburgh: Digital Curation Centre, 2013), zuletzt geprüft am 15.02.2018, http://www.dcc.ac.uk/webfm_send/1279.
- 9 Das DFG-Projekt „Grid für die Wissenschaft – WissGrid“ unter Leitung der Georg-August Universität Göttingen wurde am 30.04.2012 nach dreijähriger Laufzeit abgeschlossen, zuletzt geprüft am 15.02.2018, <https://www.sub.uni-goettingen.de/projekte-forschung/projektetails/projekt/wissgrid/>.
- 10 Jens Ludwig und Harry Enke, Hrsg., *Leitfaden zum Forschungsdaten-Management – Handreichungen aus dem WissGrid-Projekt* (Glückstadt: vwh, 2013), 83ff, zuletzt geprüft am 15.02.2018, http://www.forschungsdaten.org/images/b/b0/Leitfaden_Data-Management-WissGrid.pdf.

Die Nachnutzung von Forschungsdaten ist in den vergangenen Jahren auch in Deutschland immer stärker in den Fokus der Drittmittelgeber gerückt: Die Deutsche Forschungsgemeinschaft (DFG) verlangt seit einigen Jahren bei der Antragstellung entsprechende Konzepte zur Projektdokumentation,¹¹ hat diese auf ihrer Webseite „Umgang mit Forschungsdaten“¹² in den übergeordneten Leitlinien für Antragstellende dargelegt und die Forderungen 2015 in einer Leitlinie verankert.¹³ Das Bundesministerium für Bildung und Forschung (BMBF) macht die Erstellung von DMPs nicht generell obligatorisch für die Antragstellung, fordert allerdings in einer Reihe von Programmen – insbesondere, wenn es sich um transnationale Programme handelt – die Erstellung von DMPs.¹⁴

Das EU-Rahmenprogramm für Forschung und Innovation für die Jahre 2014–2020 „Horizon 2020“ ging einen Schritt weiter als die deutschen Förderorganisationen und machte bereits in der ersten Förderperiode ab 2014 die Erstellung eines DMPs für alle am Open Research Data (ORD) Pilot beteiligten Forschungsprojekte obligatorisch: „Das Pilotprojekt für offene Forschungsdaten betrifft sowohl die Grunddaten, die zur Reproduktion der in wissenschaftlichen Veröffentlichungen vorgestellten Ergebnisse nötig sind, als auch andere Daten, die von den Forschern in ihren Datenmanagementplänen (DMP) angegeben werden. Dies ist ein neues Element in ‚Horizont 2020‘. Ein DMP legt einen Datenverwaltungszyklus für alle Daten fest, die im Zuge des Forschungsprojekts erfasst, verarbeitet oder erzeugt werden. Alle Forschungsprojekte, die sich an dem Pilotprojekt beteiligen, müssen einen DMP haben, die anderen ‚Horizont-2020‘-Projekte sind aufgefordert, freiwillig einen solchen Plan aufzustellen, wenn dies für ihre Forschung relevant ist.“¹⁵ In der Fortschreibung des Rahmenprogramms 2017 wurde der ORD Pilot von den ursprünglich sieben thematischen Förderbereichen auf alle Themengebiete von Horizon 2020 ausgeweitet. In den aktuellen „Guidelines on FAIR Data Management in Horizon 2020“ vom Juli 2017 ist der DMP als „key element of good data management“¹⁶ definiert und erstmals die Vorlage für einen DMP, das „Horizon 2020 FAIR Data Management Plan (DMP) template“,¹⁷ als Anhang beigefügt. Während es für die Teilnehmer am ORD Pilot hinsichtlich der Veröffentlichung ihrer Forschungsdaten verschiedene opt-out-Möglichkeiten gibt,¹⁸ heißt es zur Erstellung eines DMP: „A DMP is required for all projects participating in the extended ORD pilot, unless they opt out of the ORD pilot. However, projects that opt out are still encouraged to submit

11 DFG, *Leitfaden für die Antragstellung – Projektanträge* (Bonn: DFG, 2018), 5–6, zuletzt geprüft am 15.02.2018, http://www.dfg.de/formulare/54_01/54_01_de.pdf.

12 DFG, *Umgang mit Forschungsdaten. DFG-Leitlinien zum Umgang mit Forschungsdaten*, Aktualisierungsdatum: 29.10.2015 (Bonn: DFG, 2015), zuletzt geprüft am 15.02.2018, http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/.

13 DFG, *Leitlinien zum Umgang mit Forschungsdaten*, verabschiedet durch den Senat der DFG am 30.09.2015, (Bonn: DFG, 2015), zuletzt geprüft am 15.02.2018, www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf.

14 Siehe beispielsweise: BMBF, „Bekanntmachung: Richtlinie zur Förderung transnationaler Verbundvorhaben auf dem Gebiet der Systemmedizin innerhalb des ERA-Netzes ERACoSysMed,“ *Bundesanzeiger* (03.02.2017), Punkt 2, Datenmanagement, zuletzt geprüft am 15.02.2018, <https://www.bmbf.de/foerderungen/bekanntmachung-1309.html>.

15 Siehe Pressemitteilung zum Start des Pilotprojekts: Europäische Kommission, „Kommission startet Pilotprojekt zur Öffnung der mit öffentlicher Förderung erzeugten Forschungsdaten,“ *Pressemitteilung* (16. Dezember 2013), zuletzt geprüft am 15.02.2018, http://europa.eu/rapid/press-release_IP-13-1257_de.htm.

16 European Commission, *H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020*, Version 3.0, 26. July 2016, (Brussels: European Commission, 2016), 4, zuletzt geprüft am 15.02.2018, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.

17 Ebd., 6–9.

18 Ebd., 3.

a DMP on a voluntary basis.¹⁹ Mit ihrem Förderprogramm Horizon 2020 hat die EU-Kommission seit 2014 entscheidende Anreize gesetzt, dass mit öffentlicher Förderung erzeugte Forschungsdaten frei zugänglich gemacht werden. Gleichzeitig hat sie mit der engen Kopplung von Forschungsförderung und DMP und der stetigen Argumentation, dass zusätzliche Informationen zu den Forschungsdaten für deren Nachnutzung zwingend notwendig sind, das Verständnis gefördert, dass zu jedem Forschungsprojekt ein DMP gehört.

Horizon 2020 gab auch an der TU Berlin wesentliche Impulse für die Erstellung von DMPs. In Zusammenarbeit mit Pilotpartnern wurde im Rahmen des SZF ein Web-Tool zur Erstellung von DMPs, TUB-DMP,²⁰ entwickelt und Version 0 im August 2014 für einen erweiterten Kreis zum Probetrieb freigegeben.²¹ Im September 2015 wurde Version 1 für den Routinebetrieb zur Nutzung freigeschaltet und ist seitdem an der TU Berlin im regulären Einsatz.²² Seit Dezember 2017 läuft Version 2 von TUB-DMP im Routinebetrieb, die eine Reihe von Erweiterungen enthält, wie Login mittels TUB-Account, Anlegen beliebig vieler DMPs für Teilprojekte von Verbundvorhaben, Erstellung von DMPs für Vorhaben ohne TUB-Projektnummer. TUB-DMP kann direkt oder auf den Seiten des SZF²³ in der Produktionsversion und in einer Demoversion aufgerufen werden.

1.3. Mehrwert für die Forschenden – das Konzept von TUB-DMP

Neben Beratung und Information besteht die grundlegende Zielsetzung des SZF darin, Mehrwertdienste für die Forschenden zu schaffen, um ihnen den Umgang mit Forschungsergebnissen zu vereinfachen, Arbeitsweisen zu effektivieren und die tägliche Arbeit zu erleichtern. Zur Entwicklung bedarfsgerechter, leicht bedienbarer Werkzeuge arbeitet das SZF mit Pilotpartnern aus dem Kreis der Forschenden der TU Berlin zusammen. Diesem Leitgedanken ist das SZF auch bei der Entwicklung des Web-Tools TUB-DMP gefolgt: Mit TUB-DMP wird die Erstellung eines DPMs leicht gemacht.

TUB-DMP wurde als Mehrwertdienst konzipiert und entwickelt:

- Bereits vorhandene Daten aus dem Projektantrag werden automatisch übernommen.
- Informationen zu DepositOnce, dem Repositorium der TU Berlin, sind vorab bereits eingetragen und müssen nicht mehr ermittelt werden (DepositOnce-Standardwerte).
- Der Step-by-Step-Workflow leitet die Forschenden auf einfache Weise durch einen Fragenkatalog. Er fragt alle wesentlichen Bestandteile eines DMPs ab und sorgt dafür, dass nichts Wichtiges vergessen wird.

19 Ebd., 4.

20 „TUB-DMP,“ zuletzt geprüft am 15.02.2018, <https://dmp.tu-berlin.de>.

21 Das Tool wurde erstmals auf dem Bibliothekartag 2014 in Bremen vorgestellt. Siehe Fabian Fürste, „TUB-DMP – Der Datenmanagementplan als Bindeglied zwischen Forschungsinformationssystem und Repositorium,“ (Vortrag auf dem 103. Deutschen Bibliothekartag in Bremen, 2014), Vortragsfolien unter: <https://nbn-resolving.org/urn:nbn:de:0290-opus-17540>. Siehe auch: Fabian Fürste, „TUB-DMP – Ein Werkzeug zur Erstellung von Datenmanagementplänen an der TU Berlin,“ in *DINI-Workshop „Datenmanagementpläne und ihre Bedeutung im Forschungsdatenmanagement“* (Berlin, 2015), zuletzt geprüft am 11.04.2018, <http://www.forschungsdaten.org/images/d/dc/O6--fuerst-werkzeug-erstellung-datenmanagementplaene.pdf>.

22 Monika Kuberek, „Der Umgang mit Forschungsdaten: Für ‚Horizon 2020‘-Anträge müssen Datenmanagementpläne entwickelt werden; Die Universitätsbibliothek hat ein neues Tool entwickelt, um Forschende zu unterstützen,“ *TU intern* Nr. 7-9 (2016), 4, zuletzt geprüft am 11.04.2018, <http://archiv.pressestelle.tu-berlin.de/tui/16juli/#/4>.

23 „Datenmanagementplan TUB-DMP,“ zuletzt geprüft am 11.04.2018, http://www.szf.tu-berlin.de/menue/dienste_tools/datenmanagementplan_tub_dmp.

- Im Eingabeformular werden die Daten strukturiert abgelegt und können – übersichtlich gegliedert – als Ausdruck an die Förderorganisation weitergegeben werden.
- Verschiedene Templates, so unter anderem zur Erstellung eines Horizon 2020-konformen DMPs, stehen zur Auswahl.
- In TUB-DMP können DMPs für Einzelprojekte wie auch für Verbundprojekte mit vielen Teilprojekten angelegt werden.
- Die Forschenden können sich mit ihrem TUB-Account in TUB-DMP anmelden.
- TUB-DMP verfügt über ein Helpdesk (Ticketsystem), an das sich die Forschenden bei Problemen oder Fragen wenden können. Für die Beantwortung erster Fragen zu TUB-DMP gibt es ausführliche FAQs²⁴ auf den Seiten des SZF. Das SZF-Team bietet darüber hinaus bei Bedarf Beratungen und Workshops an.

Durch die Nachnutzung der Projektantragsdaten und durch die DepositOnce-Standardwerte sind beim DCC-Template fast 30%, beim weniger umfangreichen Horizon2020-Template sogar rund 40% des DMP bereits ausgefüllt, wenn der Plan zum ersten Mal aufgerufen wird. Mit diesem konzeptionellen Ansatz des Mehrwertdienstes unterscheidet sich TUB-DMP grundlegend von dem weit verbreiteten Web-Tool DMPonline,²⁵ in dem alle Informationen manuell per Fließtext eingegeben werden müssen. Der seit November 2015 im Rahmen eines DFG-Projekts entwickelte Research Data Management Organiser (RDMO) bietet in seinem jetzigen Entwicklungsstand weder die Mehrwertdienste noch den Funktionsumfang von TUB-DMP.²⁶

TUB-DMP unterscheidet drei Kategorien von Daten und Eingabefeldern:

a) Daten, die bereits in der Forschungsprojektdatenbank der TU Berlin gespeichert sind

Um Doppelerfassung zu vermeiden und zur Arbeitserleichterung werden bereits vorhandene Projektinformationen aus der Antragstellung (wie Projekttitel, Abstract, Projektleitung, Projektlaufzeiten,...) in TUB-DMP nachgenutzt. Die Informationen werden bei der Initiierung eines Plans übertragen und liegen bereits vor, wenn eine Wissenschaftlerin oder ein Wissenschaftler ihn erstmals aufruft.

b) DepositOnce-Standardwerte

Als besonderen Mehrwert für die Forschenden, die ihre Forschungsdaten in DepositOnce speichern, gibt es die DepositOnce-Standardwerte. Das sind Kurztexte mit Informationen zu Deposit-Once (wie Infos zu den Betreibern des Repositoriums und zur Infrastruktur, Sicherheitskonzept, Speicherzyklen, usw.), die das SZF-Team vorab erstellt und in der TUB-DMP-Datenbank gespeichert hat. Auf diese Weise sind die entsprechenden Fragen in TUB-DMP bereits mit Antworten versehen, wenn

24 Nutzung von TUB-DMP, zuletzt geprüft am 11.04.2018, http://www.szf.tu-berlin.de/menue/faq/nutzung_von_tub_dmp.

25 DMPonline ist eine gemeinsame Entwicklung des Digital Curation Centre (DCC) in Großbritannien und des University of California Curation Center (UC3). Der Quellcode ist open source und auf GitHub eingestellt, zuletzt geprüft am 11.04.2018, <https://dmponline.dcc.ac.uk>.

26 Im November 2017 ist die zweite Projektförderphase von RDMO angelaufen, in deren Verlauf „... u.a. Rollenkonzepte und Anbindungen an die institutionelle Infrastruktur wie Repositorien, Ticket-Systeme und Infrastruktur für Authentifizierung und Autorisierung“ entwickelt werden sollen, wie es auf der RDMO-Website heißt, zuletzt geprüft am 11.04.2018, <https://rdmorganiser.github.io>.

Forschende ihren Plan zum ersten Mal aufrufen. Sie ersparen sich damit das Erfragen dieser Informationen bei den Betreibern von DepositOnce – eine nicht unwesentliche Arbeitserleichterung.

c) Sonstige Inhalte, die sich auf das Projekt oder die Forschungsergebnisse beziehen

Darüber hinaus ermöglicht TUB-DMP die strukturierte Eingabe weiterer Informationen in definierten Freitextfeldern.

Alle Inhalte in TUB-DMP können von den Forschenden geändert werden, einschließlich der Werte aus der Forschungsprojektdatenbank und der DepositOnce-Standardwerte, die dem spezifischen Projekt entsprechend angepasst werden können.

2. Struktur und Inhalt von TUB-DMP

In TUB-DMP können unterschiedliche Templates implementiert werden, um DMPs für spezielle Förderprogramme oder verschiedene Disziplinen zu erstellen. Als Basis-Template ist ein allgemeines Template hinterlegt, dessen Inhalt der DCC-Checklist²⁷ entspricht. Anhand des allgemeinen Templates sollen hier Struktur und Inhalt von TUB-DMP erläutert werden.

Die DCC-Checklist untergliedert sich in Sektionen: Auf die Sektion „Administrative Data“ mit allgemeinen Angaben zum Projekt folgen die Sektionen 1 bis 8, die sich auf die Forschungsdaten selbst beziehen. Entsprechend der DCC-Sektionen sind die Felder im TUB-DMP-Eingabeformular untergliedert und benannt. Das strukturierte Eingabeformular mit den vorformulierten Fragen bildet eine Art „Gerüst“ und soll den Forschenden eine Anleitung und Hilfestellung geben, welche Informationen ein DMP enthalten soll. Nicht alle Felder in TUB-DMP müssen ausgefüllt werden. Leere Felder werden beim Export ignoriert, sodass der Ausdruck – beispielsweise zur Vorlage bei einer Förderorganisation – nur die Informationen aus den ausgefüllten Feldern enthält.

Um Verbundprojekte abzubilden, verfügt TUB-DMP über eine hierarchische Struktur. Verbundvorhaben (wie DFG-Sonderforschungsbereiche, DFG-Forschergruppen, DFG-Graduiertenkollegs, BMBF- und EU-Verbundprojekte) bestehen in der Regel aus mehreren Teilprojekten, die jeweils eigene Inhalte und Zielsetzungen haben und in denen jeweils unterschiedliche Forschungsdaten generiert werden. Jedes dieser Teilprojekte benötigt einen eigenen DMP. In TUB-DMP ist es möglich, DMPs für beliebig viele Teilprojekte eines Verbundprojekts anzulegen. Dazu wurde eine hierarchische Struktur realisiert, die für Einzelprojekte wie auch für Verbundprojekte genutzt wird: Auf der obersten Ebene gibt es den „Projekt-DMP“,²⁸ darunter pro Teilprojekt einen eigenen „DMP“. Zu einem „Projekt-DMP“ können beliebig viele „DMPs“ angelegt werden. Das Verhältnis „Projekt-DMP“:„DMP“ ist bei einem Einzelprojekt 1:1, bei einem Verbundvorhaben 1:n.

²⁷ DCC-Checklist.

²⁸ Die Bezeichnungen „Projekt-DMP“ und „DMP“ sind eventuell etwas verwirrend und seien hier kurz erläutert: Bei der Implementierung der hierarchischen Struktur mussten unterschiedliche Bezeichnungen für die beiden Hierarchiestufen gewählt werden; die Entscheidung fiel auf „Projekt-DMP“ für die oberste und „DMP“ für die darunterliegende Hierarchiestufe. Diese technischen Begriffe und ihre Unterscheidung sind nur im Zusammenhang mit der hierarchischen Struktur von Bedeutung und hier in Anführungszeichen gesetzt. Ansonsten ist in diesem Artikel mit der Abkürzung DMP immer die Gesamtheit des Plans gemeint, der sich aus „Projekt-DMP“ und „DMP“ zusammensetzt.

Der „Projekt-DMP“ enthält die übergeordneten Projektinformationen (wie Titel, Abstract, Projektleitung usw.) entsprechend der DCC-Sektion „Administrative Data“. Der „DMP“ enthält die Informationen zu den Forschungsdaten, entsprechend der DCC-Sektionen 1–8. Die Inhalte des „Projekt-DMP“ werden bei der Initiierung des Plans aus der Forschungsprojektdatenbank übernommen. Im „DMP“ finden sich dann die vorab generierten DepositOnce-Standardwerte und weitere Textfelder, die von den Wissenschaftlerinnen und Wissenschaftlern manuell ausgefüllt werden.

3. Der Workflow in TUB-DMP

Das Schema in Abbildung 2 zeigt den Workflow in TUB-DMP – inklusive der Akteure (Wissenschaftlerinnen und Wissenschaftler sowie SZF-Team) auf der linken Seite und den Datenquellen / Schnittstellen zu TUBIS²⁹ (Authentifizierung) sowie zur Forschungsprojektdatenbank (Übernahme der Projektdaten) auf der rechten Seite.

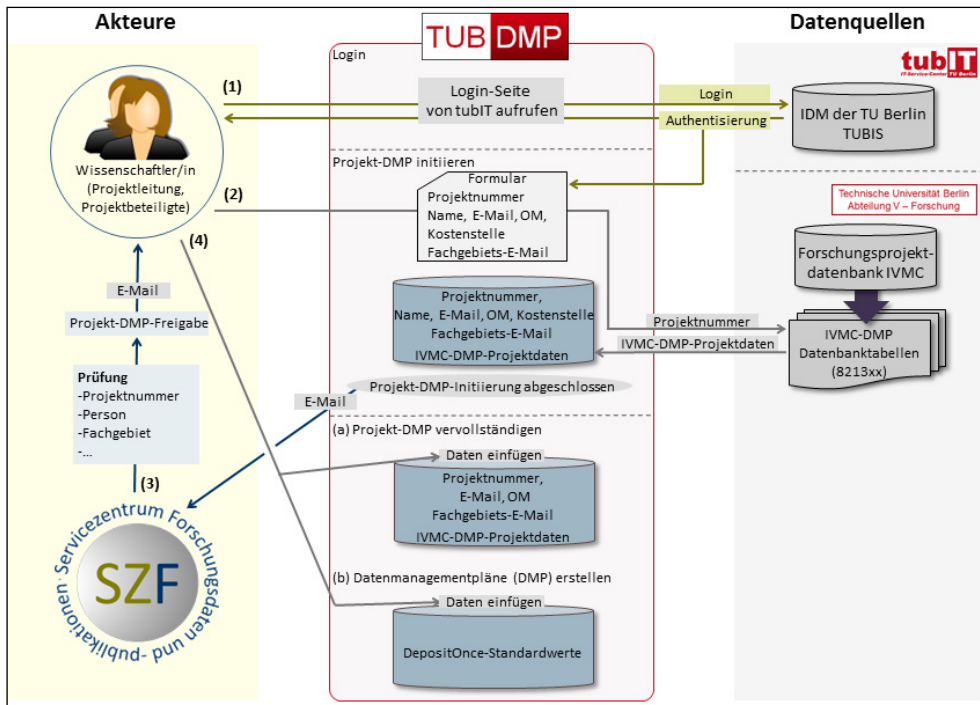


Abb 2: Schematische Darstellung des Workflows in TUB-DMP

Im Folgenden werden die Schritte im Einzelnen beschrieben; das Verfahren ist für Einzelprojekte und für Verbundprojekte mit vielen Teilprojekten identisch:

²⁹ TUBIS (TU Berlin Integrations Service) ist das von tubIT betriebene, rollenbasierte Identitätsmanagement der TU Berlin.

a) Authentifizierung (Wissenschaftler/innen)

Beim Anklicken des Login-Buttons auf der TUB-DMP-Homepage werden die Wissenschaftlerinnen und Wissenschaftler über zur Loginseite von tubIT weitergeleitet und authentifizieren sich dort mit ihrem TUB-Account.

b) Plan initiieren (Wissenschaftler/innen)

Nach dem Login erhalten sie in TUB-DMP ein Formular mit einigen Authentifizierungsmerkmalen aus TUBIS (wie Name, Fachgebiet), das sie mit der TUB-Projektnummer (so vorhanden) und der institutionellen E-Mail-Adresse des Fachgebiets vervollständigen. Sobald das Formular abgeschickt wird, wird ein „Projekt-DMP“ generiert.

- Ist das Feld mit der Projektnummer ausgefüllt und war die Anfrage in den IVMC-DMP-Datenbanktabellen positiv, werden die Projektdaten in die TUB-DMP-Datenbank eingefügt und stehen dann den Wissenschaftlerinnen und Wissenschaftlern in den entsprechenden Feldern des „Projekt-DMP“ zur Bearbeitung zur Verfügung.
- Ist das Feld mit der Projektnummer ausgefüllt und war die Anfrage in den IVMC-DMP-Datenbanktabellen negativ, wird ein leerer „Projekt-DMP“ mit der angegebenen Projektnummer generiert.
- Ist das Feld mit der Projektnummer leer, wird eine zufällig generierte Projektnummer vom System erzeugt und in der TUB-DMP-Datenbank gespeichert.

Beim Abschicken des Formulars wird zudem das SZF-Team in der UB per E-Mail darüber informiert, dass ein neuer „Projekt-DMP“ angelegt wurde.

c) Formalprüfung (UB)

Bevor der Plan von den Wissenschaftlerinnen und Wissenschaftlern bearbeitet werden kann, erfolgt eine formale Prüfung durch das SZF-Team, ob seine Initiierung plausibel war (anhand der Einträge im „Projekt-DMP“, Zugehörigkeit der eintragenden Person zum Fachgebiet usw.). Ist die Plausibilität gegeben, wird der Plan vom SZF-Team freigegeben und eine E-Mail generiert, mit der die Person, die den Plan initiiert hat, von der Freigabe in Kenntnis gesetzt wird.

d) Vervollständigung der Daten (Wissenschaftler/innen)

Nach Freigabe des „Projekt-DMP“ können die Wissenschaftlerinnen und Wissenschaftler diesen, falls notwendig, vervollständigen und einen „DMP“ mit den projektspezifischen Informationen anlegen.

Alle Änderungen werden in der TUB-DMP-Datenbank gespeichert. Zur weiteren Verwendung können die Wissenschaftlerinnen und Wissenschaftler den Datensatz als PDF exportieren.

Falls Änderungen vorgenommen werden sollen, kann ein DMP über das Login jederzeit wieder aufgerufen und bearbeitet werden.

4. Technische Umsetzung

4.1. Das Web-Tool TUB-DMP

TUB-DMP ist eine Eigenentwicklung der UB. Sie übernimmt im Rahmen des SZF auch die Pflege des Web-Tools.

a) Software

Das Web-Tool setzt auf dem modernen PHP-Framework Laravel auf, welches wiederum auf dem Symfony-Framework basiert. TUB-DMP wird über den Apache-Webserver ausgeliefert. Die DMPs werden in einer relationalen Datenbank gespeichert und verwaltet (TUB-DMP-Datenbank); als Datenbanksoftware wird PostgreSQL verwendet. Die Anwendung nutzt weitverbreitete Tech Stacks und stellt keine besonderen Anforderungen an Hard- und Software. Der Quellcode von TUB-DMP wird auf GitHub als Open Source Software veröffentlicht.³⁰

b) Integration in die IT-Infrastruktur der TU Berlin

Die schematische Darstellung in Abbildung 3 zeigt die Integration von TUB-DMP in die IT-Infrastruktur der TU Berlin, die vom Rechenzentrum der TU Berlin, dem IT-Service-Center tubIT, betrieben wird. Auf diese Weise sind zum einen eine hohe Ausfallsicherheit sowie ein hoher Schutz der Anwendung und der Daten gewährleistet, zum anderen werden durch die Nachnutzung der vorhandenen Infrastruktur Synergien genutzt, die den Betriebsaufwand für TUB-DMP verringern.

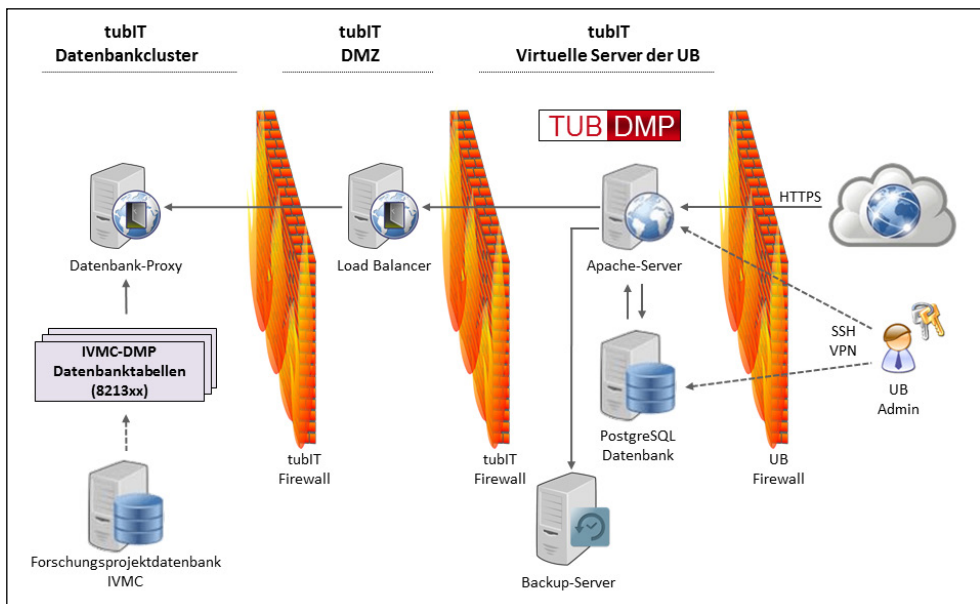


Abb 3: Integration von TUB-DMP in die IT-Infrastruktur der TU Berlin

30 In Vorbereitung. GitHub-Account der UB der TU Berlin, zuletzt geprüft am 11.04.2018, <https://github.com/tuub>.

Die Datenbank und das Dateisystem von TUB-DMP laufen auf virtuellen Servern der UB, die Teil der von tubIT betriebenen virtualisierten Serverumgebung der TU Berlin sind. Die in TUB-DMP eingegebenen Daten werden im Backup-System der TU Berlin, das ebenfalls von tubIT betrieben wird, gesichert. Der Zugriff von außen auf TUB-DMP ist durch verschlüsselte Netzwerkverbindungen geschützt: Nur die TUB-DMP-Administratorinnen und -Administratoren der UB haben im Campusnetz der TU Berlin – mittels SSH (Secure Shell) in Verbindung mit VPN (Virtual Private Network) – Zugriff auf die Datenbank und das Dateisystem. Für den Web-Zugang der Benutzerinnen und Benutzer auf TUB-DMP ist das sichere Kommunikationsprotokoll HTTPS (Hypertext Transfer Protocol Secure) implementiert.

Der Abruf der Projektdaten aus der Forschungsprojektdateiabank, um sie in TUB-DMP zu übernehmen, erfolgt über eine spezielle Kommunikationsschnittstelle, einen sogenannten Proxy. Dieser, wie auch die Forschungsprojektdateiabank, die sensible Daten enthält, liegen in einem besonders sicheren Bereich von tubIT – geschützt durch eine DMZ (Demilitarized Zone). Um zu verhindern, dass die Forschungsdatenbank durch eine zu große Anzahl von Anfragen beeinträchtigt wird, ist ein sogenannter Load-Balancer zwischengeschaltet. Aus Datenschutzgründen erfolgt von TUB-DMP kein direkter Zugriff auf die Forschungsprojektdateiabank, sondern lediglich auf einen eigens für TUB-DMP definierten Auszug, der in den IVMC-DMP-Datenbanktabellen abgelegt ist. Das Verfahren zur Übernahme der Projektdaten und weitere Erläuterungen zur Skizze sind im nachfolgenden Kapitel beschrieben.

4.2. Automatisierte Übernahme von Projektdaten nach TUB-DMP

Die Forschungsprojektdateiabank IVMC (Integriertes Verwaltungs- und Management Controlling) ist die zentrale Datenbank der TU Berlin, in der die Informationen zu ihren Forschungsprojekten gespeichert werden. Für den DMP werden nicht die Daten aus der ursprünglichen Projektantragstellung verwendet, sondern die Daten, die nach der Projektbewilligung von der Forschungsabteilung (Abt. V) aktualisiert, gegebenenfalls korrigiert oder auch ergänzt wurden. Die Übernahme der Projektdaten aus IVMC nach TUB-DMP wurde in Zusammenarbeit der drei Kooperationspartner im SZF – UB, tubIT und Abt. V – konzipiert und umgesetzt. Eindeutiger Identifikator ist die TUB-Projektnummer, die in der Forschungsprojektdateiabank hinterlegt ist.

Aus Datenschutzgründen darf es keinen direkten Zugriff von TUB-DMP auf die IVMC-Datenbank geben. Daher wurden von UB und Abt. V die für die Nachnutzung relevanten Felder definiert – eine kleine, eng gefasste und fest definierte Untermenge der Forschungsprojektdateiabank mit allgemeinen Projektinformationen. Zur Datenübernahme werden diese Daten in separaten Datenbanktabellen (IVMC-DMP-Datenbanktabellen) auf dem IVMC-Datenbankserver abgelegt und in einem spezifischen Nummernkreis (8213xx) bereitgestellt. Für die Transaktion ist die UB mit einem Service-Account gegenüber dem IVMC-Datenbankserver authentifiziert. Während des Anfrage-Vorgangs ist die Verbindung zwischen dem TUB-DMP-Server und der IVMC-DMP-Datenbank durch eine Transportverschlüsselung gesichert.

4.3. Funktionen von TUB-DMP

a) Zugriffsschutz und Authentifizierung über Shibboleth

Das Konzept von TUB-DMP sieht vor, dass die Forschenden ihre DMPs in einem zugriffsgeschützten Bereich erstellen und bearbeiten. In Zusammenarbeit mit tubIT hat die UB in TUB-DMP Shibboleth implementiert, so dass sich die Wissenschaftlerinnen und Wissenschaftler mittels ihres TUB-Accounts authentifizieren können.³¹

b) Nutzung unterschiedlicher Templates

Es werden Templates für verschiedene Zwecke zur Verfügung gestellt, die per Mausklick aus einer Liste ausgewählt werden können. Zurzeit ist neben dem allgemeinen Template ein Template zur Erstellung Horizon 2020-konformer DMPs implementiert. Es basiert auf dem „Horizon 2020 FAIR Data Management Plan (DMP) template.“³² Weitere Templates sind in Vorbereitung.

c) Speicherfunktion und neue Versionen

Jeder DMP soll bei Bedarf, dem Projektfortgang entsprechend, geändert oder aktualisiert werden. Damit wird den Anforderungen von Horizon 2020 entsprochen, wo vorgesehen ist, dass der DMP im Laufe des Projektfortschritts weitergeschrieben und aktualisiert wird.³³

Während der Eingabe in TUB-DMP kann die Arbeit jederzeit unterbrochen und der Plan zwischengespeichert werden. Die Einträge bleiben erhalten und der Plan kann beim erneuten Aufruf weiterbearbeitet werden. Beim Abspeichern wird jeweils das Datum der letzten Änderung automatisch erstellt und gespeichert. Ist die Erstellung eines Plans abgeschlossen, kann er mit Klick auf einen spezifischen Button als Version abgespeichert werden; danach ist die Bearbeitung nicht mehr möglich. Soll der DMP geändert werden, kann beim nächsten Aufruf eine neue Version angelegt werden. Diese enthält die Daten der Vorgängerversion, die geändert und bearbeitet werden können.

Bei der Ersterstellung eines DMP wird das Erstellungsdatum automatisch generiert und gespeichert; es bleibt als konstantes Datum immer erhalten. Im DMP werden das Datum der Ersterstellung und das Datum der letzten Änderung angezeigt. Für die bearbeitende Person sind in TUB-DMP die verschiedenen Versionen sichtbar und können als PDF exportiert, jedoch nicht mehr verändert werden.

d) Exportfunktion

Ein DMP kann als PDF exportiert werden, um beispielsweise einem Projektbericht beigelegt zu werden. Beim Export werden nur die Felder, die ausgefüllt wurden, berücksichtigt und nur deren Inhalte im PDF angezeigt. Leer gelassene Felder werden beim Export übergangen, so dass keine „Lücken“ im Text entstehen.

31 tubIT stellt einen Shibboleth-Dienst zur verteilten Authentisierung und Autorisierung für Webanwendungen und Webservices bereit: Ein Benutzer muss sich nur einmal bei seiner Heimateinrichtung (hier: TU Berlin) authentisieren, um ortsunabhängig auf Dienste oder lizenzierte Inhalte verschiedener Anbieter (hier: TUB-DMP) zugreifen zu können (Single-Sign-on).

32 European Commission, H2020, 6–9.

33 Ebd., 5.

5. Perspektive

Eines der Prinzipien des SZF ist die bedarfsbezogene Entwicklung und Weiterentwicklung benutzerorientierter Anwendungen, so auch bei TUB-DMP. Die Zusammenarbeit mit Pilotpartnern wird fortgeführt und TUB-DMP entsprechend der Anforderungen der Forschenden weiterentwickelt. Im Fokus steht derzeit die Implementierung weiterer Templates, sowohl auf die Anforderungen von Förderorganisationen bezogen als auch disziplinspezifische Templates. Auch weitere Exportmöglichkeiten (csv, XML) sind vorbereitet und können bei Bedarf zügig implementiert werden.

In Planung ist die Verlinkung zwischen TUB-DMP und DepositOnce, um die DMPs zu veröffentlichen. TUB-DMP ist inhaltlich eng an das Repository DepositOnce gekoppelt: In DepositOnce werden Ergebnisse aus Forschungsprojekten, in TUB-DMP die DMPs mit Informationen zu den zugehörigen Forschungsprojekten gespeichert. Die Veröffentlichung soll durch Verlinkung des DMPs in DepositOnce realisiert werden, ist somit im eigentlichen Sinne keine Funktion von TUB-DMP, sondern von DepositOnce, sei hier im Zusammenhang mit dem DMP aber trotzdem erwähnt. Die Wissenschaftlerinnen und Wissenschaftler sollen selbst entscheiden, ob sie ihre DMPs veröffentlichen, und dies auf einfache Weise tun können, indem sie die DMPs beim jeweiligen Forschungsergebnis in DepositOnce verlinken.

Literaturverzeichnis

- BMBF. „Bekanntmachung: Richtlinie zur Förderung transnationaler Verbundvorhaben auf dem Gebiet der Systemmedizin innerhalb des ERA-Netzes ERACoSysMed.“ *Bundesanzeiger*, 03.02.2017. Zuletzt geprüft am 15.02.2018. <https://www.bmbf.de/foerderungen/bekanntmachung-1309.html>.
- DFG. *Leitfaden für die Antragstellung: Projektanträge*. Bonn: DFG, 2018. Zuletzt geprüft am 15.02.2018. http://www.dfg.de/formulare/54_01/54_01_de.pdf.
- DFG. *Leitlinien zum Umgang mit Forschungsdaten*. Bonn: DFG, 2015. Zuletzt geprüft am 15.02.2018. http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf.
- DFG. *Umgang mit Forschungsdaten: DFG-Leitlinien zum Umgang mit Forschungsdaten*. Aktualisierungsdatum: 29.10.2015. Bonn: DFG, 2018. Zuletzt geprüft am 15.02.2018. http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/.
- Digital Curation Centre. *Checklist for a Data Management Plan: v4.0*. Edinburgh: Digital Curation Centre, 2013. Zuletzt geprüft am 15.02.2018. http://www.dcc.ac.uk/webfm_send/1279.

- Donnelly, Martin und Sarah Jones. *Data Management Plan Content Checklist: Draft Template for Consultation*. Edinburgh: Digital Curation Centre, 2009. Zuletzt geprüft am 15.02.2018. http://www.dcc.ac.uk/webfm_send/1280.
- European Commission. *H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020*. Version 3.0, 26. July 2016. Brussels: European Commission, 2016. Zuletzt geprüft am 15.02.2018. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.
- Europäische Kommission. „Kommission startet Pilotprojekt zur Öffnung der mit öffentlicher Förderung erzeugten Forschungsdaten.“ *Pressemitteilung* (16. Dezember 2013). Zuletzt geprüft am 15.02.2018. http://europa.eu/rapid/press-release_IP-13-1257_de.htm.
- Fürste, Fabian. „TUB-DMP – Der Datenmanagementplan als Bindeglied zwischen Forschungsinformationssystem und Repositorium“. Vortrag auf dem 103. Deutschen Bibliothekartag in Bremen, 2014. Vortragsfolien unter <https://nbn-resolving.org/urn:nbn:de:0290-opus-17540>.
- Fürste, Fabian. „TUB-DMP – Ein Werkzeug zur Erstellung von Datenmanagementplänen an der TU Berlin.“ *DINI-Workshop „Datenmanagementpläne und ihre Bedeutung im Forschungsdatenmanagement“*. Berlin, 2015. Zuletzt geprüft am 11.04.2018. <http://www.forschungsdaten.org/images/d/dc/06--fuerst--werkzeug-erstellung-datenmanagementplaene.pdf>.
- Kuberek, Monika. *Organisatorisch-technisches Konzept für eine Forschungsdaten-Infrastruktur in der TU Berlin*. Berlin: Technische Universität, 2012. <http://dx.doi.org/10.14279/depositonce-6604>.
- Kuberek, Monika. „Der Umgang mit Forschungsdaten: Für ‚Horizon 2020‘-Anträge müssen Datenmanagementpläne entwickelt werden; Die Universitätsbibliothek hat ein neues Tool entwickelt, um Forschende zu unterstützen.“ *TU intern* Nr. 7-9 (2016), 4. Zuletzt geprüft am 13.04.2018. <http://archiv.pressestelle.tu-berlin.de/tui/16juli/#/4>.
- Ludwig, Jens und Harry Enke, Hrsg. *Leitfaden zum Forschungsdaten-Management – Handreichungen aus dem WissGrid-Projekt*. Glückstadt: vwh, 2013. Zuletzt geprüft am 15.02.2018. http://www.forschungsdaten.org/images/b/b0/Leitfaden_Data-Management-WissGrid.pdf.
- Wilkinson, Mark D. et al. „The FAIR Guiding Principles for scientific data management and stewardship.“ *Sci. Data* 3:160018 (2016). <https://doi.org/10.1038/sdata.2016.18>.

Offene Forschungsdaten an der Universität Heidelberg: von generischen institutionellen Repositorien zu fach- und projektspezifischen Diensten

Jochen Apel, Universitätsbibliothek Heidelberg

Fabian Gebhart, Universitätsrechenzentrum Heidelberg

Leonhard Maylein, Universitätsbibliothek Heidelberg

Martin Wlotzka, Universitätsrechenzentrum Heidelberg

Zusammenfassung:

Die Universität Heidelberg hat 2014 das Kompetenzzentrum Forschungsdaten als gemeinsame Serviceeinrichtung der Universitätsbibliothek und des Universitätsrechenzentrums eingerichtet. Der vorliegende Beitrag stellt die Angebote des Kompetenzzentrums zur Publikation von Forschungsdaten vor, fasst bisherige Erfahrungen zusammen und diskutiert auf dieser Grundlage exemplarisch die Rolle von institutionellen Veröffentlichungsplattformen für Open Research Data. Im Einzelnen werden dabei das institutionelle Datenrepositorium heiDATA, die Bild- und Multimediadatenbank heiICON sowie die derzeitige Weiterentwicklung des Dienstleistungsportfolios des Kompetenzzentrums im Rahmen des Projekts „Community-spezifische Forschungsdatenpublikation (CS-FDP)“ vorgestellt.

Summary:

In 2014 Heidelberg University established the Competence Centre for Research Data as a joint facility of the University Library and the university's Computing Centre. This article describes the Competence Centre's services for publishing research data and examines on that basis the role of institutional publication platforms for open research data. In particular the paper discusses the institutional research data repository heiDATA, the image and multimedia database heiICON and the current refinement of the Competence Centre's service portfolio within the project "Community Specific Research Data Publication (CS-FDP)".

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S61-71>

Autorenidentifikation: Apel, Jochen: GND 1025737180, ORCID: <https://orcid.org/0000-0002-0395-4120>; Gebhart, Fabian: GND 1148481745, ORCID: <https://orcid.org/0000-0003-2466-8120>; Maylein, Leonhard: GND 1019546514, ORCID: <https://orcid.org/0000-0002-9423-489X>; Wlotzka, Martin: GND 1031120645, ORCID: <https://orcid.org/0000-0003-1794-4799>

Schlagwörter: Forschungsdaten; Forschungsdatenmanagement; Forschungsdatenrepositorium; Open Research Data

1. Forschungsdaten- und Publikationsservices an der Universität Heidelberg

Das Kompetenzzentrum Forschungsdaten der Universität Heidelberg (KFD) ist eine gemeinsame Serviceeinrichtung der Universitätsbibliothek und des Universitätsrechenzentrums, die Wissenschaftlerinnen und Wissenschaftler beim strukturierten und nachhaltigen Management ihrer Forschungsdaten

unterstützt.¹ Der vorliegende Beitrag stellt die Angebote des KFD zur Publikation von Forschungsdaten vor, fasst bisherige Erfahrungen zusammen und diskutiert auf dieser Grundlage exemplarisch die Rolle von institutionellen Veröffentlichungsplattformen für Open Research Data. Im Einzelnen werden dabei das institutionelle Datenrepositorium heiDATA², die Bild- und Multimediadatenbank heidICON³ sowie die derzeitige Weiterentwicklung der KFD-Dienstleistungen im Rahmen des MWK-geförderten Projekts Community-spezifische Forschungsdatenpublikation (CS-FDP)⁴ vorgestellt.

Diese Dienste zur Forschungsdatenpublikation sind zum einen in das umfassendere Serviceportfolio des KFD zum Forschungsdatenmanagement (FDM) eingebettet. Das KFD bietet darüber hinaus fachliche Beratung zu technischen und organisatorischen Themen sowie Informationen über rechtliche Aspekte. Dies umfasst auch Unterstützung bei der Erstellung von Projektanträgen und Datenmanagementplänen. Hinzu kommen generische IT-Services zur aktiven Arbeit mit Daten im Projekt, z.B. Dateiablage und -Sharing mit dem Dienst heiBOX, die zentrale Speicherung wissenschaftlicher Daten durch den Service SDS@HD oder der Zugang zu Ressourcen für das Hochleistungsrechnen.⁵ Zum anderen ergänzen die Datenpublikationsdienste die breite Palette von universitären Angeboten zur Open-Access-Veröffentlichung, die derzeit folgende Publikationsmöglichkeiten umfasst:⁶

- heiUP: Heidelberg University Publishing
- heiBOOKS: Heidelberger E-Books
- heiJOURNALS: Heidelberger Open-Access-Zeitschriften
- heiDOK: Heidelberger Dokumentenserver
- heidICON: Heidelberger Bild- und Multimediadatenbank
- heiDATA: Heidelberger Forschungsdatenrepositorium
- Publikationsplattformen für Altertumswissenschaften, Kunstgeschichte und Asienwissenschaft im Kontext der Fachinformationsdienste (FID)
- Digitale Editionen für Textcorpora

2. heiDATA – das Forschungsdatenrepositorium der Universität Heidelberg

Mit der Einrichtung des KFD wurde 2014 zugleich die Einführung eines institutionellen Forschungsdatenrepositoriums beschlossen. Das Repositorium soll insbesondere für diejenigen Forschenden der Universität ein Angebot zur Datenpublikation zur Verfügung stellen, in deren Forschungsbereichen es keine geeigneten fachlichen Repositorien gibt. Darüber hinaus soll es eine Möglichkeit zur Veröffentlichung für Datensupplemente schaffen, die Textpublikationen auf den oben genannten Heidelberger Open-Access-Plattformen ergänzen. heiDATA, so der Name der Plattform, wurde im Sommer 2014

1 Kompetenzzentrum Forschungsdaten der Universität Heidelberg, <http://data.uni-heidelberg.de/>.

2 heiDATA, <https://heidata.uni-heidelberg.de/>.

3 heidICON, <http://heidicon.ub.uni-heidelberg.de>.

4 CS-FDP, <https://www.urz.uni-heidelberg.de/de/cs-fdp>.

5 Vgl. „Dienste und Angebote.“ KFD, zuletzt geprüft am 22.05.2018, <http://data.uni-heidelberg.de/dienste.html>.

6 Eine Übersicht über diese Plattformen findet sich hier: „Elektronisches Publizieren – Open Access,“ Universität Heidelberg, zuletzt geprüft am 22.05.2018, www.openaccess.uni-hd.de.

auf Grundlage der Open-Source-Software Dataverse⁷ implementiert und bietet folgende wesentliche Eigenschaften, die für nachhaltige Datenpublikationen erforderlich sind:

- Persistente zitierfähige Adressierung von Datensätzen über DOI
- Unterstützung unterschiedlicher Open-Content-Lizenzen
- Verknüpfung von Forschungsdaten mit zugehörigen Publikationen
- Individuelle Berechtigungsmodelle für Zugriffe (inkl. Zugriffsmöglichkeiten für Gutachter vor der Freischaltung via Private URLs)
- Publikation unterschiedlicher Dateiformate
- Unterstützung verschiedener domänenspezifischer Metadatenstandards (z.B. DDI u.a. für sozialwissenschaftliche Daten, ISA-Tab für die Lebenswissenschaften oder das VOResource Schema für die Astronomie sowie weitere)⁸
- Untergliederung in Sektionen, sog. Dataverses, z.B. für Projekte, Forschungsgruppen, Institute, Journals oder auch einzelne Personen
- Kuration der Daten auf Datensatz- und Dateiebene als Dienstleistung des KFD
- Nachweis aller Datensätze im Data-Citation-Index⁹ sowie weiteren Nachweissystemen über offene API
- Nachweis von Datenpublikationen in der Hochschulbibliographie heiBIB basierend auf der Katalogisierung der Datenpublikationen im Südwestdeutschen Bibliotheksverbund (möglich für eigene und externe Repositorien durch Verlinkung)¹⁰

Die Dataverse-Software wird maßgeblich am Harvard Institute for Quantitative Social Science (IQSS) entwickelt. Weltweit gibt es bislang 33 öffentliche Instanzen der Software (Stand: 23.4.2018). Deren größte ist das Harvard Dataverse mit über 75.000 publizierten Datensätzen, das als generisches Angebot Forschenden weltweit die Möglichkeit bietet, kostenfrei Forschungsdaten zu publizieren.¹¹ Interessant ist zudem das DataverseNL, das als gemeinsames Angebot von 13 niederländischen Forschungseinrichtungen betrieben wird.¹² Die Entwicklung von Dataverse erfolgt als Open-Source-Projekt. Die Community vernetzt sich, neben der offenen Softwareentwicklung und dem zugehörigen Austausch auf GitHub¹³, über ein Forum¹⁴, monatliche Community Calls¹⁵ sowie jährliche Community-Meetings.¹⁶

7 The Dataverse project, <https://dataverse.org/>.

8 DDI steht für den gleichnamigen Metadatenstandard der Data Documentation Initiative, ISA-Tab ist das sog. Investigation/Study/Assay tab-delimited Format für Metadaten aus dem Bereich der omics-basierten Methoden, das VOResource Schema ist das Metadatenchema der International Virtual Observatory Alliance. Eine Übersicht über sämtliche von Dataverse unterstützten Metadatenstandards inkl. Verlinkungen zu deren Dokumentationen findet sich auf <http://guides.dataverse.org/en/latest/user/appendix.html>.

9 Der Data Citation Index ist ein separat lizenzierbarer Teilindex des Web of Science, der ausgewählte Forschungsdatenpublikationen sowie Zitierungen dieser Datensätze nachweist. „Data Citation Index,“ Clarivate Analytics, zuletzt geprüft am 22.05.2018, http://wokinfo.com/products_tools/multidisciplinary/dci/

10 heiBIB, <http://heibib.uni-hd.de>.

11 Harvard Dataverse, <https://dataverse.harvard.edu/>

12 DataverseNL, <https://dataverse.nl/>

13 <https://github.com/IQSS/dataverse>.

14 <https://dataverse.org/forum>.

15 <https://dataverse.org/community-calls>.

16 <https://dataverse.org/events>.

Wesentliches Charakteristikum des in Heidelberg verfolgten Ansatzes zur institutionellen Datenpublikation ist die Übernahme von Kurationsarbeiten durch das KFD. Ziel ist es hierbei, Datensätze bereits während des Ingest-Prozesses so für das Repositorium aufzubereiten, dass sie soweit als möglich in einer Form vorliegen, die eine Langzeitarchivierung der Daten erlaubt. Dies umfasst die Überprüfung und die Validierung von Dateiformaten unter Einsatz gängiger Tools wie DROID und JHOVE, wobei man jedoch konstatieren muss, dass ein sinnvoller Einsatz dieser Systeme nur für eine verhältnismäßig geringe Zahl von Dateiformaten möglich ist.¹⁷ Unter Umständen werden Datenbestände vor der Publikation in archivfähige Formate konvertiert.¹⁸ Auch die Verbesserung der fachwissenschaftlichen Dokumentation des Datensatzes durch geeignete Metadaten und kontrollierte Vokabulare wird durch das KFD in enger Abstimmung mit den Forschenden unterstützt.

2.1. Erfahrungen nach drei Jahren heiDATA - lohnt sich der Aufwand?

Nach mittlerweile über drei Jahren Produktivbetrieb lässt sich zumindest ein erstes Zwischenfazit für den Betrieb von heiDATA ziehen, anhand dessen sich die grundsätzliche Frage nach der Sinnhaftigkeit von institutionellen Veröffentlichungsplattformen für Forschungsdaten zumindest in Ansätzen diskutieren lässt. Eine naheliegende Haltung hierzu ist, dass es sich bei Forschungsdaten – im Unterschied zu Textpublikationen – um deutlich heterogeneres und interpretationsbedürftigeres Material handelt. Und damit könnten sie zum einen nur durch Fachleute aus der jeweiligen Disziplin angemessen kuratiert werden und zum anderen nur in spezifisch zugeschnittenen Fachrepositorien wohldokumentiert und auffindbar publiziert werden. Diese Allgemeinplätze der Forschungsdatendiskussion sind sicher zutreffend.¹⁹ Auch die bisherige, eher verhaltene Nutzung von heiDATA legt dies nahe. Bislang wurden auf der Plattform 105 Datensätze veröffentlicht (Stand: 23.4.2018). Vor dem Hintergrund, dass die Universität Heidelberg eine der forschungs- und publikationsstärksten Universitäten Europas mit aktuell ca. 5.800 Wissenschaftlerinnen und Wissenschaftlern ist, ist dies auf den ersten Blick keine besonders hohe Zahl. Zieht man den Vergleich zu institutionellen Datenpublikationsplattformen weiterer deutscher Universitäten, so stellt man fest, dass sich dort im Großen und Ganzen ein ähnliches Bild ergibt.²⁰

17 DROID, <https://digital-preservation.github.io/droid/>; JHOVE, <http://jhove.sourceforge.net/>.

18 Beispielhaft sei hier auf einen Datensatz mit archäologischen Grabungsdaten verwiesen, bei dessen Publikation ca. 1.100 Bilddateien vom proprietären CorelDraw-Format in Vektorgrafiken konvertiert wurden (unter wiederholter Rückkopplung mit dem verantwortlichen Wissenschaftler, um Informationsverlust durch die Konversion möglichst ausschließen zu können): Paul A. Yule, „Pottery Drawings, Zafar, Jemen, Mostly Excavated,“ *heiDATA*, V3, 6. April 2017, <https://doi.org/10.11588/data/10068>.

19 Vgl. hierzu auch Louise Corti u. a., *Managing and Sharing Research Data: A Guide to Good Practice* (Los Angeles, Calif. [u.a.]: SAGE, 2014), 197–201.

Nebenbei bemerkt: In der Praxis muss man sich die vermeintlich geeigneteren fachlichen Repositorien freilich immer im Detail anschauen. Für große, lange und gute etablierte Dienste gilt dies sicherlich, aber recherchiert man z.B. im Kontext der Beratung von Forschenden quer durch re3data (<https://www.re3data.org/>), so stößt man auf etliche Fachrepositorien, die weder in Bezug auf Datendokumentation und Metadatenstandards besondere fachbezogene Tiefe bieten noch im Hinblick auf die nachhaltige Verfügbarkeit des Materials Vertrauen erwecken.

20 Beispielhaft sei hier auf die folgenden universitären Dienste verwiesen, die entweder eigenständige Forschungsdatenrepositorien oder kombinierte Repositorien für Texte, Forschungsdaten und weitere Materialien sind. Bei den kombinierten Repositorien kann man sich in der Regel über die Browserfunktionalität den Veröffentlichungstyp „Forschungsdaten“ oder verwandte Kategorien anzeigen lassen: LMU München, Open Data LMU, <https://data.ub.uni-muenchen.de/>; TU Berlin, DepositOnce, <https://depositonce.tu-berlin.de/>; Universität Bielefeld, PUB, <https://pub.uni-bielefeld.de/data>; Universität Freiburg, FreIDok plus, <https://freidok.uni-freiburg.de/>; Universität Mannheim, MADATA, <https://madata.bib.uni-mannheim.de/>; Universität Tübingen, FDAT, <https://fdat.escience.uni-tuebingen.de/portal/#/start>; MPG, EDMOND, <https://edmond.mpdl.mpg.de/imeji/>.

Was aber heißt dies? Sind und bleiben generische institutionelle Datenrepositorien ein Nischenangebot, das mittelfristig sogar wieder komplett verschwinden wird? Tatsächlich ist das Publizieren von Forschungsdaten in vielen Disziplinen noch Neuland und diejenigen Disziplinen, in denen dies nicht der Fall ist, verfügen über geeignete und gut genutzte Fachrepositorien. Umgekehrt sind es gerade diejenigen Disziplinen ohne „Open-Data-Kultur“ in denen es noch keine fachspezifischen Publikationsplattformen für Forschungsdaten gibt, aber in denen doch zunehmend ein Wissenschaftlerinteresse an offenen Forschungsdaten entsteht. Hier füllen institutionelle Repositorien eine Leerstelle. Dies macht ein Blick auf die Verteilung der auf heiDATA publizierten Datensätze deutlich: Die Veröffentlichungen stammen aus neun der zwölf Fakultäten der Universität und aus vier zentralen wissenschaftlichen Einrichtungen und verteilen sich somit breit über die an der Universität Heidelberg vertretenen Disziplinen. Hinzu kommen Datensupplemente zu Textveröffentlichungen der Heidelberger Publikationsdienste. Veröffentlicht werden dabei ganz unterschiedliche Datentypen von biophysikalischen Simulationsdaten oder Umfragedaten aus dem Bereich der Public Health über Textcorpora aus der Computerlinguistik bis hin zu archäologischen Grabungsdaten oder Ergebnissen von Distant-Reading-Analysen theologischer Texte. Mehr als 9.000 Downloads seit dem Beginn des Betriebs sprechen dabei zudem dafür, dass die publizierten Datensätze durchaus in breiterem Ausmaß von Dritten genutzt oder zumindest gesichtet werden. Selbst wenn also die Nachfrage rein quantitativ (noch?) nicht exorbitant hoch ist, so kann man dennoch feststellen, dass sie sich breit über verschiedene Disziplinen verteilt. Genau an diesen sog. „long tail“ richtet sich entsprechend das generische Angebot institutioneller Forschungsdatenrepositorien.²¹

Eine hieran anschließende Frage, die sich angesichts der oben dargestellten Erfahrungen stellt, ist, ob sich der Aufwand lohnt, eigene institutionelle Angebote zur Forschungsdatenveröffentlichung zu schaffen. Die Heidelberger Antwort auf diese Frage lautet: Unser Ziel ist es, dauerhaft Verantwortung für die Forschungsdaten der eigenen Institution zu übernehmen. Eine Datenpublikation auf heiDATA soll mehr sein als das bloße Abladen von Dateien auf einem Server und das Registrieren einer DOI für diese Dateien. Sie sollen vielmehr den FAIR-Data-Principles genügen.²² Wenn es sich tatsächlich als richtig erweisen sollte, dass Forschungsdaten über sämtliche oder zumindest hinreichend viele Wissenschaftsdisziplinen hinweg so wertvolle Ressourcen sind, dass sie langzeitarchiviert und möglichst breit zugänglich gemacht werden sollten, dann müssen Betreiber von Forschungsdatenrepositorien notwendigerweise die Kuration der Daten und deren Langzeitarchivierung als ihre Aufgabe sehen. Bei zukünftig zunehmendem Bedarf müssen diese Aufgaben dann innerhalb der forschungsinfrastrukturellen Arbeitsteilung auf viele Schultern verteilt werden – in dieser Weise skizziert es auch der Rat für Informationsinfrastrukturen in seinen Überlegungen zur „Nationalen Forschungsdateninfrastruktur“.²³

21 Zur Rolle des „long tail“ im Kontext wissenschaftlicher Daten vgl. P. Bryan Heidorn, „Shedding Light on the Dark Data in the Long Tail of Science,“ *Library Trends* 57, Nr. 2 (2008): 280–299, <https://doi.org/10.1353/lib.0.0036>; Jillian C. Wallis, Elizabeth Rolando und Christine L. Borgman, „If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology,“ *PLOS ONE* 8, Nr. 7 (Juli 2013): e67332, <https://doi.org/10.1371/journal.pone.0067332>.

22 Mark D. Wilkinson u. a., „The FAIR Guiding Principles for Scientific Data Management and Stewardship,“ *Scientific Data*, 15. März 2016, <https://doi.org/10.1038/sdata.2016.18>.

23 Vgl. RfII – Rat für Informationsinfrastrukturen, *Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland* (Göttingen, 2016), zuletzt geprüft am 22.05.2018, <http://www.rfii.de/?wpdmdl=1998>.

Wie diese Arbeitsteilung im Detail aussehen wird, ist dabei weitgehend offen. Es ist keineswegs ausgemacht, dass es über alle Disziplinen hinweg fachliche Datenarchive geben wird (ebenso wenig wie es zum aktuellen Zeitpunkt überhaupt ausgemacht ist, ob es den Bedarf dafür wirklich übergreifend gibt). Im Gegenteil, die Wahrscheinlichkeit, dass hierbei auch institutionelle Dienste eine wichtige Rolle spielen werden, ist durchaus hoch. Dies gilt umso mehr, als man auch im Hinblick auf die Nachhaltigkeit verschiedener fachlicher Repositorien durchaus kritisch sein darf. Insbesondere in den Fächern, für die das Publizieren von Forschungsdaten noch neu ist, werden aktuell auf nationaler und internationaler Ebene mit erfreulich hoher Drittmittelunterstützung etliche Datenrepositorien eingerichtet. Welche von diesen sich aber tatsächlich im Fach etablieren und den Sprung in einen nachhaltigen Regelbetrieb inklusiver gesicherter Finanzierung schaffen werden, ist unklar.

Weiterhin gibt es auch von Seiten der Forschenden durchaus Nachfrage nach institutionellen Angeboten. Lokale FDM-Serviceeinheiten sind nah an den Wissenschaftlerinnen und Wissenschaftlern, Absprachen und persönliche Treffen sind kurzfristig und mit wenig Aufwand zu realisieren, man kennt sich aus anderen Kontexten und arbeitet daher vertrauensvoll zusammen, kleinere und mittlere Datenmengen können ggf. kostenfrei publiziert werden etc. Lokale Repositorien bieten den Datengebern mit geringem Aufwand ein Schaufenster für die Datenprodukte des eigenen Instituts, der eigenen Forschungsgruppe oder anderer Teileinheiten; insbesondere die oben dargestellte Untergliederung von heiDATA in unterschiedliche Dataverses ist hierbei hilfreich.²⁴

2.2. Repositorium oder Repositorien?

heiDATA ist ein eigenständiges Forschungsdatenrepositorium, das vor Ort parallel zu vier EPrints-Instanzen (universitärer Dokumentenserver heiDOK plus drei Fachrepositorien für die FIDs), dem mit der Software easydb betriebenen Bild- und Multimediarepositorium heidICON sowie weiteren OJS- und OMP-basierten Publikationsplattformen betrieben wird. Damit stellt sich auch die Frage, ob der Betrieb einer solchen Repositorienlandschaft nicht deutlich mehr Aufwand erzeugt als ein gemeinsames Repositorium für Texte, Forschungsdaten und ggf. weitere Materialien. Und tatsächlich sind entsprechende Lösungen verbreitet, teilweise mit Integration weiterer Funktionalitäten für Universitätsbibliographien oder Forschungsinformationssysteme. Beispielhaft sei hier auf die Systeme PUB an der Uni Bielefeld²⁵, DepositOnce der TU Berlin²⁶ oder FreiDok plus an der Universität Freiburg²⁷ verwiesen.

Dass heiDATA als eigenständiges System betrieben wird, liegt primär in einer pragmatischen Einschätzung begründet: Andere Lösungen wären im Hinblick auf Administration und Betrieb (vermutlich) deutlich arbeitsintensiver gewesen. Nachdem 2014 die Entscheidung zum Angebot eines Services zur Datenpublikation gefallen war, wurden hierfür die Systeme EPrints, DSpace, ESciDoc und Dataverse gesichtet. Eine zentrale Anforderung an die Software war dabei die Möglichkeit der schnellen Implementierung eines Services mit wenig Anpassungsbedarf, z.B. beim Datenschema. Zum damaligen

24 Beispiele: <https://heidata.uni-heidelberg.de/dataverse/awiexeco> oder <https://heidata.uni-heidelberg.de/dataverse/asiaeurope>.

25 <https://pub.uni-bielefeld.de/>

26 <https://depositonce.tu-berlin.de/>

27 <https://freidok.uni-freiburg.de/>

Zeitpunkt erfüllte Dataverse insbesondere dieses Kriterium am besten. Darüber hinaus überzeugt Dataverse im laufenden Betrieb bislang auch durch umfassende Funktionalitäten, eine aufgeräumte und strukturierte Benutzeroberfläche, wenig administrativen Aufwand und stabilen Betrieb sowie eine aktive, schnelle und an den Wünschen der Nutzercommunity ausgerichtete Entwicklungsarbeit.

Der tatsächliche Arbeitsaufwand für die Erstinstallation lässt sich dabei grob mit ca. zwei Personenwochen beziffern, der in 2017 durchgeführte größere Versionssprung auf Dataverse Version 4.x inklusive Migration der publizierten Datenbestände mit ca. zwei bis drei Wochen. Die für den laufenden Betrieb anfallenden systemadministrativen Aufgaben haben aber einen sehr geringen Umfang. Der Hauptaufwand entsteht im Kontext der Beratung der Forschenden und im Zuge der Kurationsaktivitäten beim Daten-Ingest. Und dieser Aufwand entsteht ohnehin plattformunabhängig und auch unabhängig davon, ob man ein eigenständiges Forschungsdatenrepositorium oder ein integriertes System betreibt.

2.3. Und wenn es sich anders entwickelt? Exitstrategie

Wie im bisherigen Verlauf des Beitrags beschrieben, ist nicht klar, wie sich das Thema Open Research Data abseits der ohnehin datenpublizierenden Disziplinen entwickeln wird. Interesse und breite Akzeptanz auf Seiten der Wissenschaftler müssen sich noch entwickeln, und auch ob tatsächlich in substantiellem Maß eine Nachnutzung von publizierten Forschungsdaten stattfindet, welche die Versprechen des „Data Web“²⁸ und der allgegenwärtigen Data-Science- und Big-Data-Rhetorik einlöst, muss sich noch erweisen. Damit einhergehend ist auch offen, wie sich die Landschaft der fachlichen und institutionellen Datenrepositorien zukünftig entwickeln wird.

Die Heidelberger Dataverse-Instanz wurde und wird daher als Übergangslösung betrachtet, deren Eignung vor dem Hintergrund des sich entwickelnden Felds des Forschungsdatenmanagements kontinuierlich neu bewertet werden muss. Sämtliche organisatorischen und technischen Festlegungen sind so gestaltet, dass die Datenbestände perspektivisch in andere Systeme, seien es übergreifende Dienste oder neue institutionelle Lösungen, überführt werden können. Diese Überlegung war ebenfalls ein wesentliches Kriterium bei der Entscheidung für Dataverse als Softwarebasis für heiDATA: Die Anwendung ist kostenfrei, open source und vor allem ist der Export von Daten und Metadaten in strukturierter Form problemlos möglich, sodass bei einer entsprechenden Neubewertung die Überführung des in heiDATA publizierten Materials in fachspezifische Dienste oder übergreifende Lösungen für Institutionen wie RADAR oder KITOpenData (vormals bwDataDiss) prinzipiell möglich ist.²⁹

28 Michael Nielsen, *Reinventing Discovery: The New Era of Networked Science* (Princeton: Princeton University Press, 2012), 110–116.

29 RADAR und KITOpenData sind zentrale, nicht-kommerzielle Repositoriendienste, die von Institutionen kostenpflichtig in Anspruch genommen werden können, um ihren Forschenden eine Möglichkeit zur Datenpublikation zu bieten. Vgl. RADAR, <https://www.radar-service.eu>; KITOpenData, <https://bwdatadiss.kit.edu/>.

3. heidICON oder wie zählt man eigentlich Datenpublikationen?

Für ein vollständiges Bild der Heidelberger Publikationsangebote für Forschungsdaten lohnt es sich zusätzlich zu heiDATA auch das seit 2005 an der Universität Heidelberg betriebene Bild- und Multimediaarchiv heidICON in den Blick zu nehmen.³⁰ heidICON basiert auf der Software easyDB, einem Media-Asset-Management-System der Firma Programmfabrik und bietet umfangreiche Funktionalitäten zur Verwaltung, Bereitstellung und Erschließung von AV-Daten in offenen und geschlossenen Medienpools.³¹ Für die Erschließung wird dabei das normierte Vokabular der GND sowie das Metadatenschema LIDO eingesetzt.³²

In heidICON sind gegenwärtig über 510.000 Objekte abgelegt (Stand: 23.4.2018), die im Schnitt auf ca. 260.000 Einzelobjektaufrufe pro Jahr kommen. Dies ist im Kontext dieses Aufsatzes ein interessantes Faktum, da die in heidICON publizierten AV-Materialien gemäß der einschlägigen Definitionen selbstverständlich auch unter das Lemma „Forschungsdaten“ fallen und man somit mit guten Gründen sagen kann, dass in Heidelberg (an vielen anderen Universitäten verhält es sich ähnlich) seit 2005 ein Forschungsdatenrepositorium betrieben wird, in dem bereits über eine halbe Million Forschungsdatensätze publiziert wurden. Setzt man dies zu den oben erwähnten 105 Datensätzen auf heiDATA in Beziehung, wird eine weitere Schwierigkeit bei der quantitativen Beurteilung von Forschungsdatenrepositorien deutlich: Es ist eben nicht ohne weiteres klar, wie man Forschungsdaten zählt. Was ist das Verhältnis von Einzeldateien (und den ihnen zugehörigen Metadaten) zu einem Gesamtdatensatz (und ihm zugehörigen Metadaten)? Wann bildet eine Menge von Dateien gemeinsam einen Datensatz, wann eher eine Kollektion oder einen Pool von Datensätzen? Usw.

4. CS-FDP – ein Projekt zur Entwicklung fach- und communityspezifischer Datenportale

Insbesondere für umfangreichere Datensammlungen ist das Angebot einer individuell zugeschnittenen Präsentationsoberfläche eine Anforderung, die häufig von Seiten der Forschenden formuliert wird. heiDATA bietet dies nicht, sondern stellt letztendlich nur eine Downloadoption für die Daten bereit. Gleichzeitig gibt es in vielen Fällen für solche Datensammlungen keine geeigneten Fachrepositorien, die eine ansprechende Präsentation der Daten ermöglichen würden. Dies führt dazu, dass Forschungsprojekte vielfach projekteigene Datenbanken mit Webpräsentation aufsetzen, deren dauerhafter Betrieb über die Projektlaufzeit hinaus jedoch nicht gesichert ist. Diese Beobachtung bildet die Ausgangslage für die gegenwärtige Weiterentwicklung der Forschungsdatenservices des KFD, die im Rahmen des Projekts „Community-spezifische Forschungsdatenpublikation“ (CS-FDP) vorangetrieben wird.³³ CS-FDP wird gefördert vom baden-württembergischen Ministerium für Wissenschaft, Forschung und Kunst (MWK) und ist gemeinsam mit weiteren Projekten aus den Bereichen

30 heidICON, <http://heidicon.ub.uni-heidelberg.de>.

31 Programmfabrik, <https://www.programmfabrik.de/>

32 „LIDO,“ ICOM International Committee for Documentation, zuletzt geprüft am 22.05.2018, <http://www.lido-schema.org/>.

33 „Community-spezifische Forschungsdatenpublikation,“ Universitätsrechenzentrum Heidelberg, zuletzt geprüft am 22.05.2018, <https://www.urz.uni-heidelberg.de/de/cs-fdp>.

Forschungsdatenmanagement und Virtuelle Forschungsumgebungen Teil der E-Science-Initiative des Landes Baden-Württemberg.³⁴

Ziel von CS-FDP ist der Aufbau eines überschaubaren Pools von generischen, durch das KFD bereitgestellten Softwarewerkzeugen zur Erstellung dynamischer Forschungsdatenportale, mit denen eine Vielzahl der bei unterschiedlichen Projekten auftretende Anforderungen erfüllt werden können, die sich aber gleichzeitig mit vertretbarem Aufwand zentral administrieren und archivieren lassen. Darüber hinaus ist es Ziel von CS-FDP, ein tragfähiges Konzept für die Langzeitarchivierung der universitären Forschungsdatenbestände zu entwickeln. CS-FDP soll damit die Lücke zwischen generischer, universitärer FDM-Infrastruktur und individuellen, fachcommunity-spezifischen Anforderungen verkleinern. Das KFD kooperiert mit fünf Pilotpartnern, mit denen unterschiedliche technische Lösungen prototypisch umgesetzt und auf ihre Eignung für ein breiteres Serviceangebot geprüft werden.

Im ersten Pilotprojekt wird in Zusammenarbeit mit dem Heidelberger Institut für Semiotik eine zukunftsfähige Lösung für das am Institut mit DFG-Unterstützung aufgebaute „Semitische Tonarchiv“ erarbeitet, das für die semiotische Forschung ein zentrales Angebot an Audiodaten bereitstellt.³⁵ Die dem bisherigen Institutsarchiv zugrundeliegende Software ist mittlerweile stark veraltet und eine adäquate technische Betreuung kann von Seiten des Instituts nicht mehr gewährleistet werden. Im Rahmen von CS-FDP werden die ca. 3.000 dort archivierten Tondokumente nach heidICON überführt. Ziel ist es, heidICON als Backend für die Daten zu verwenden und darauf aufsetzend eine eigenständige Rechercheoberfläche anzubieten, die unabhängig von heidICON ist. Diese Oberfläche soll spezifische Darstellungs- und Rechercheoptionen realisieren, die die semiotische Forschungscommunity zur Arbeit mit den Daten benötigt (z.B. Geovisualisierung und -browsing). Realisiert wird das Frontend mit der Software Vue.js unter Nutzung der easyDB-API.³⁶

Im Rahmen des zweiten Pilotprojekts wird ebenfalls die Software easyDB eingesetzt, allerdings für eine reine Metadatenverwaltung (ohne direkten Bezug zu AV-Medien). Mit easyDB können komplexe Datenmodelle schnell, flexibel und ohne Programmieraufwand abgebildet werden. Ein ausgefeiltes Rechte- und Tag-Management ist dabei inbegriffen. Ein Anwendungsfall hierfür wird gemeinsam mit dem DFG-Projekt „Dynastinnen und Bettelorden im spätmittelalterlichen Reich. Weibliche Frömmigkeit zwischen Hof, Stadt und Kloster (1250–1400)“, kurz „Gender & Piety“ erprobt. Hier geht es darum, komplexe soziale und räumliche Beziehungen in historischen Netzwerken abzubilden.

Im Rahmen des dritten Pilotprojekts „Digitale Papyrussammlung“ wird der Aufbau individueller, webbasierter Rechercheoberflächen mit Hilfe der Software Django³⁷ sowie deren Verknüpfung mit standardisierten Metadaten - in diesem Fall Publikationsmetadaten gemäß des TEI/EpiDoc-Standards

34 Allgemeine Informationen zur E-Science-Initiative: „E-Science: Wissenschaft unter neuen Rahmenbedingungen“, Ministerium für Wissenschaft, Forschung und Kunst, zuletzt geprüft am 22.05.2018, <https://mwk.baden-wuerttemberg.de/de/forschung/forschungslandschaft/e-science/>; Projektübersicht: <https://www.forschungsdaten.info/praxis-kompakt/aktuelle-infrastrukturprojekte/>.

35 www.semarch.uni-hd.de/.

36 Vue.js, <https://vuejs.org/>.

37 Django, <https://www.djangoproject.com/>.

– entwickelt.³⁸ Die Präsentation der digitalen Faksimiles sowie der Transkriptionen erfolgt über DWork, dem von der Universitätsbibliothek Heidelberg entwickelten Workflowmanagement- und Präsentationssystem für Retrodigitalisate, dessen Editionsmodul zu diesem Zweck erweitert wurde.³⁹ Insbesondere die Visualisierung und Normalisierungen von editorischen Eingriffen unter Berücksichtigung von fachspezifischen Eigenheiten wurden überarbeitet und deutlich stärker konfigurierbar gestaltet (sowohl für den Editierenden als auch für den Leser). Diese Entwicklungen können bereits jetzt im Editionsprojekt „Documenta Nepalica“ der Heidelberger Akademie der Wissenschaften nachgenutzt werden, in dem ein umfangreiches Korpus von Dokumenten und Texten zur Religions- und Rechtsgeschichte des vormodernen Nepal digital bereitgestellt wird.⁴⁰

Für ein viertes, lebenswissenschaftliches Pilotprojekt zur Zellbiologie wurde die freie Mikroskopiesoftware OMERO in der vom Universitätsrechenzentrum betriebenen Cloudinfrastruktur heiCLOUD eingerichtet.⁴¹ Derzeit wird sie vom Institut für Anatomie und Zellbiologie für die Analyse und das Management von Mikroskopiedaten genutzt. Zuletzt erfolgte die Anbindung großer Speicherressourcen auf der Large Scale Data Facility mit Hilfe des Dienstes SDS@hd.⁴² Dies ist ein entscheidender Schritt, um die anfallenden Datenmengen im Multi-Terabyte-Bereich in OMERO importieren, analysieren und publizieren zu können. Im weiteren Verlauf des Projektes ist geplant, eine Präsentationsoberfläche für die Publikation von Mikroskopiedaten basierend auf OMERO zu erstellen.

Der Forschungsbereich „Biostratigraphie und Paläoökologie“ am Institut für Geowissenschaften betreibt derzeit eine lokale Instanz der Software Specify.⁴³ Specify dient der Erfassung, Taxierung und digitalen Präsentation geologischer und biologischer Sammlungen. Eine zentral gehostete Instanz wurde in der heiCLOUD installiert.⁴⁴ Derzeit wird im Rahmen des fünften Pilotprojekts die Migration der lokalen Instanz des Instituts für Geowissenschaften in die zentrale Instanz vorbereitet. Anhand der gewonnenen Erfahrungen aus dieser Pilotanwendung soll das Portfolio um Specify-Unterstützung für weitere Anwender erweitert werden.

5. Zusammenfassung

Die Universität Heidelberg treibt den Ausbau universitärer Infrastrukturen für offene Forschungsdaten voran. Das Kompetenzzentrum Forschungsdaten als zuständige Serviceeinrichtung betreibt bereits seit drei Jahren mit heiDATA ein institutionelles Forschungsdatenrepositorium auf der Basis der Software Dataverse. Hinzu kommt das bereits seit 2005 in Betrieb befindliche Repositorium heiICON für AV-Daten sowie ein im Rahmen des MWK-geförderten Projekts CS-FDP im Aufbau

38 „EpiDoc: Epigraphic Documents in TEI XML.“ Sourceforge, zuletzt geprüft am 22.05.2018, <https://sourceforge.net/p/epidoc/wiki/Home/>.

39 DWork – Heidelberger Digitalisierungsworkflow, <http://www.ub.uni-heidelberg.de/helios/digi/dwork.html>.

40 „Religions- und rechtsgeschichtliche Quellen des vormodernen Nepal.“ Heidelberger Akademie der Wissenschaften, zuletzt geprüft am 22.05.2018, http://www.haw.uni-heidelberg.de/forschung/forschungsstellen/nepal/publ_docs.de.html.

41 OMERO, <https://www.openmicroscopy.org/omero/>.

42 SDS@hd – Scientific Data Storage, <https://www.urz.uni-heidelberg.de/de/sds-hd>.

43 Specify, <http://www.sustain.specifysoftware.org/>.

44 heiCLOUD, <https://heicloud.uni-heidelberg.de/heiCLOUD>.

befindlicher Service für die Erstellung und Betreuung individueller Forschungsdatenportale. Nachfrage und Nutzung der Angebote dokumentieren zwar eine zunehmende Bedeutung des Themas Open Research Data sowie die Relevanz institutioneller Datenpublikationsplattformen, zeigen aber gleichzeitig die unterschiedliche und sich im Fluss befindende Ausgangssituation in verschiedenen Fachdisziplinen auf, sodass die aufgebaute technische Infrastruktur kontinuierlich auf ihre Eignung und Bedarfsangemessenheit geprüft werden muss.

Literaturverzeichnis

- Corti, Louise, Veerle Van den Eynden, Libby Bishop und Matthew Woollard. *Managing and Sharing Research Data: A Guide to Good Practice*. Los Angeles, Calif. [u.a.]: SAGE, 2014.
- Heidorn, P. Bryan. „Shedding Light on the Dark Data in the Long Tail of Science.“ *Library Trends* 57, Nr. 2 (2008): 280–299. <https://doi.org/10.1353/lib.0.0036>.
- Nielsen, Michael. *Reinventing Discovery: The New Era of Networked Science*. Princeton: Princeton University Press, 2012.
- Rfll – Rat für Informationsinfrastrukturen. *Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland*. Göttingen, 2016. Zuletzt geprüft am 22.05.2018. <http://www.rfii.de/?wpdmdl=1998>.
- Wallis, Jillian C., Elizabeth Rolando und Christine L. Borgman. „If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology“l. *PLOS ONE* 8, Nr. 7 (Juli 2013): e67332. <https://doi.org/10.1371/journal.pone.0067332>.
- Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, u. a. „The FAIR Guiding Principles for Scientific Data Management and Stewardship.“ *Scientific Data*, 15. März 2016. <https://doi.org/10.1038/sdata.2016.18>.
- Yule, Paul A. „Pottery Drawings, Zafar, Jemen, Mostly Excavated.“ *heiDATA*, V3, 6. April 2017. <https://doi.org/10.11588/data/10068>.

Pflichtablieferung von Dissertationen mit Forschungsdaten an die DNB

Anlagerungsformen und Datenmodell

Dirk Weisbrod, Deutsche Nationalbibliothek (bis Jan. 2018, seitdem DIPF) Frankfurt am Main

Zusammenfassung:

Im Rahmen des DFG-Projektes „Elektronische Dissertationen Plus“ (eDissPlus) entwickeln die Humboldt-Universität zu Berlin (HU) und die Deutsche Nationalbibliothek (DNB) Lösungen für eine zeitgemäße Archivierung und Publikation von Forschungsdaten, die im Zusammenhang mit Promotionsvorhaben entstehen. Dabei müssen die unterschiedlichen Anlagerungsformen von Forschungsdaten an eine Dissertation berücksichtigt und in einem Datenmodell abgebildet sowie das von der DNB verwendete Metadatenschema XMetaDissPlus überarbeitet werden. Das ist notwendig, um die Relationen zwischen der Dissertation und den abgelieferten Forschungsdaten-Supplementen sowie den Daten, die auf externen Repositorien verbleiben sollen, nachzuweisen und im Katalog der DNB recherchierbar zu machen. Dieser Beitrag stellt das Datenmodell und die Änderungen im Metadatenschema vor.

Summary:

As a part of the DFG founded project “Electronic dissertations plus” (eDissPlus), the Humboldt University of Berlin (HU) and the German National Library (DNB) develop a prototype for an integrated system for archiving and publishing the research data generated or used by doctoral students as part of their dissertation project. In doing so, the different connections between thesis and research data must be considered and represented in a data model. Also, the metadata format XMetaDissPlus has to be revised according to the data model. This is necessary in order to trace the relationships between the thesis and the research data (which can either be submitted as a supplement to the document or be stored in an external repository) and make them searchable in DNB’s catalog. The paper describes the data model and the changes made in XMetaDissPlus.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S72-78>

Autorentifizikation: Weisbrod, Dirk: GND: 1079147012, ORCID: <http://orcid.org/0000-0002-9455-4527>

Schlagwörter: Forschungsdaten; Dissertationen; Anlagerungsformen; Datenmodell; Metadaten

1. Voraussetzungen für die Pflichtablieferung von Forschungsdaten an die DNB

Die Pflichtablieferungsverordnung (PflAV) der DNB definiert die Rahmenbedingungen, nach denen Forschungsdaten an die DNB abgeliefert werden können. Demnach sollen „Elemente, Software und Werkzeuge [gesammelt werden], die in physischer oder in elektronischer Form erkennbar zu den ablieferungspflichtigen Netzpublikationen gehören“ (§ 7, Absatz 2 PflAV), während selbständig veröffentlichte Primär-, Forschungs- und Rohdaten ausdrücklich nicht unter die Ablieferungspflicht

fallen (§ 9 PflAV).¹ Forschungsdaten, die erkennbar mit einer ablieferungspflichtigen Netzpublikation in Verbindung stehen, können somit von der DNB gesammelt, langzeitarchiviert und in ihrem Katalog nachgewiesen werden.² Um dieses Szenario beispielhaft durchzuspielen, erarbeitete die DNB im DFG-Projekt eDissPlus zusammen mit der Humboldt-Universität zu Berlin Lösungen für Forschungsdaten, die im Rahmen von Dissertationsprojekten an der HU generiert und zusammen mit der Dissertation abgeliefert werden. Ein erstes Ergebnis des Projektes ist die im vergangenen Jahr veröffentlichte Policy für dissertationsbezogene Forschungsdaten der DNB.³ Die HU entwickelte zudem einen Ingest-Prozess für die Publikation von dissertationsbezogenen Forschungsdaten auf ihrem Publikationsserver. In einem weiteren Schritt musste der Pflichtablieferungsworkflow der DNB für die Ablieferung von Forschungsdaten angepasst werden. Es musste folglich ein Datenmodell entwickelt werden, das den Bezug zwischen Dissertation und Forschungsdaten abbildet und eine angemessene Weiterbearbeitung der Forschungsdaten und der dazugehörigen Metadaten bei Pflichtablieferung und Langzeitarchivierung ermöglicht. Über diesen Teilaspekt von eDissPlus berichtet der folgende Artikel.

2. Von den Anlagerungsformen zum Datenmodell

Die erkennbare Zugehörigkeit von Forschungsdaten zu einer Netzpublikation ergibt sich unter anderem aus der Anlagerungsform und damit aus der formalen Beziehung zwischen Forschungsdaten und Publikation bei der Ablieferung. Dieser Aspekt war für die Entwicklung des Datenmodells besonders wichtig, da die Verarbeitungsmöglichkeiten der Daten von deren Ablieferung abhängen. Die Data Publication Pyramid (Abbildung 1), die im Rahmen des EU-Projektes Opportunities für Data Exchange (ODE) unter Teilnahme der DNB erarbeitet wurde, visualisiert diese Anlagerungsformen, indem sie Forschungsdaten in Bezug auf ihre Aufbereitung im Forschungs- und Publikationskontext in 5 Typen einteilt.⁴

Für eDissPlus war allerdings nur die Spitze der Pyramide relevant, da reine Datenpublikationen (Typ 4) sowie nicht publizierte Daten (Typ 5) keinen erkennbaren Konnex zu einer ablieferungspflichtigen Netzpublikation besitzen und deswegen von der DNB nicht berücksichtigt werden dürfen. Außerdem konnten bei der Entwicklung des Datenmodells die vollständig in eine Dissertationsschrift integrierten Daten vernachlässigt werden (Typ 1). Das sind zum Beispiel Grafiken oder Tabellen, die in ein PDF eingebettet oder in einer Printversion abgedruckt werden. Forschungsdaten, die in dieser Form vorliegen, sammelt die DNB schon immer – sozusagen automatisch als Bestandteil der abgabepflichtigen Hochschulschrift. Diese Daten werden nicht gesondert verzeichnet. Relevant für

1 Vgl. „Verordnung über die Pflichtablieferung von Medienwerken an die Deutsche Nationalbibliothek,“ zuletzt geprüft am 01.01.2018, <http://www.gesetze-im-internet.de/pflav/index.html>.

2 Vgl. hierzu Dirk Weisbrod, *Policy der Deutschen Nationalbibliothek für dissertationsbezogene Forschungsdaten*, Version 1.1 (Frankfurt, M.: Deutsche Nationalbibliothek, 2017), [urn:nbn:de:101-2017092701_3](https://nbn-resolving.org/urn:nbn:de:101-2017092701_3); „Die DNB fordert die Ablieferung von Forschungsdaten nicht aktiv ein, sondern überlässt die Definition entsprechender Rahmenbedingungen der jeweiligen Hochschule.“

3 Vgl. Weisbrod, *Policy der Deutschen Nationalbibliothek*.

4 Vgl. Susan Reilly et al., *Opportunities of Data Exchange: Report on Integration of Data and Publications* (2011), 5–6, zuletzt geprüft am 01.01.2018, <http://hdl.handle.net/10013/epic.40198>.

die Lösung der oben erläuterten Fragestellung sind hingegen die Anlagerungsformen „Supplemente“ (Typ 2) und „referenzierte Daten“ (Typ 3).

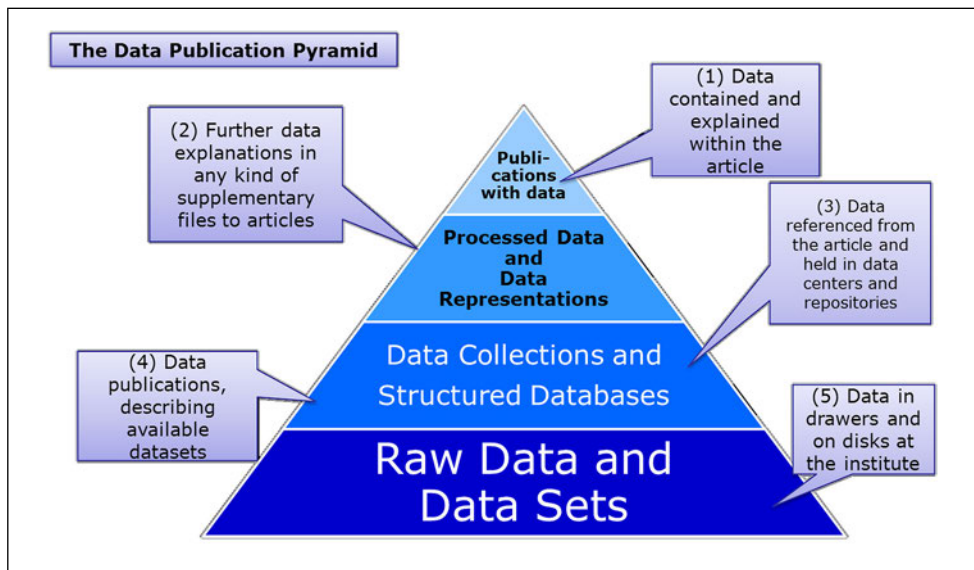


Abb. 1: Data Publication Pyramid⁵

Forschungsdaten-Supplemente konnten schon in der Vergangenheit, etwa auf einer CD-ROM oder einem anderen Datenträger, einer gedruckten Dissertation beigelegt werden. Im Falle einer Netzpublikation kann man sich vorstellen, dass Promovierende Forschungsdaten-Supplemente zusammen mit der Dissertation auf den Publikationsserver einer Hochschule hochladen und von dort die Dissertation und die zugehörigen Supplemente an die DNB abgeliefert werden. Bislang wurden Supplemente allerdings nicht im Katalog der DNB nachgewiesen. Mit eDissPlus soll sich das nun ändern.

Referenzierte Daten sind Forschungsdaten, die zwar in einer Dissertation verwendet, jedoch nicht mit dieser abgeliefert werden. So kann es vorkommen, dass Promovierende ein renommiertes Fachrepositorium oder ein institutseigenes Datenarchiv nutzen und bei der Ablieferung lediglich den Persistenten Identifikator (PID) der Daten (z.B. DOI, URN oder handle) angeben. Der Integrationsgrad zwischen Publikation und Daten ist hier geringer als bei Supplementen, da die Forschungsdaten zunächst nicht physisch an die DNB abgeliefert werden. Man kann aber davon ausgehen, dass auch referenzierte Daten, ebenso wie die Supplemente, wesentlich zum Verständnis und zur Nachprüfbarkeit der in der Dissertation präsentierten Forschungsergebnisse beitragen. Im Rahmen von eDissPlus wurde deshalb entschieden, dass die PIDs der Daten wenigstens in den Metadaten der abzuliefernden Dissertation verzeichnet und damit dauerhaft referenziert werden. Für beide Anlagerungsformen musste nunmehr ein Datenmodell für die Pflichtablieferung gefunden werden.

⁵ Vgl. Reilly et al, *Opportunities of Data Exchange*, 6.

3. Forschungsdaten-Supplemente

Ausgehend vom existierenden Workflow bot sich für die Projektbeteiligten zunächst eine einfache Variante für die Ablieferung von Supplementen an: Dissertationsschrift und Forschungsdaten in einem Container. Da die DNB für die Ablieferung von Dissertationen in der Regel eine OAI-PMH-Schnittstelle nutzt, versehen die Abliefernden oder ein entsprechend beauftragter Dienstleister (im Fall von Dissertationen also die Hochschulbibliotheken) einen solchen Container mit einer Transfer-URL und stellen diese an der Schnittstelle zu Verfügung. Die DNB harvestet dann den Container und speist diesen in ihren Pflichtablieferungs-Workflow ein. Dieses Vorgehen ist aber sehr unbefriedigend, da nunmehr der gesamte Container, also Dissertation und Forschungsdaten gemeinsam, durch die mitgelieferten Metadaten beschrieben werden. Das hat gleich zwei Nachteile. Erstens können dann die Forschungsdaten aufgrund fehlender Metadaten nicht mehr eigenständig im Katalog der DNB verzeichnet werden; andererseits können Dissertation und Daten im gesamten Pflichtablieferungs- und Langzeitarchivierungs-Workflow nicht mehr ohne weiteres voneinander unterschieden werden, ein forschungsdatenspezifisches Preservation Planning ist dann nicht mehr möglich.

In eDissPlus wurde stattdessen ein Datenmodell erarbeitet, das Dissertation und Forschungsdaten schon bei der Ablieferung deutlich voneinander trennt. Jede Entität erhält vom Abliefernden einen eigenen Metadatensatz im DNB-eigenen Format für Hochschulschriften XMetaDissPlus. Zugleich wird in diesem Modell auch der Bezug zwischen Dissertation und Daten abgebildet, denn der Nachweis dieses Bezugs ist ja Voraussetzung dafür, dass die Daten überhaupt von der DNB gesammelt werden dürfen. Somit können zukünftig Forschungsdaten mit eigenem Metadaten-Satz an der OAI-PMH-Schnittstelle bereitgestellt werden. Dabei handelt es sich um einen im tar- oder zip-Format gepackten Ordner, in dem die abgelieferten Supplemente beliebig angeordnet werden können. Auch die Ablieferung mehrerer solcher Forschungsdaten-Ordner ist möglich, wenn Promovierende oder die abliefernde Hochschule das für sinnvoll halten und damit zum Beispiel Teilprojekte der Dissertation oder wichtige Einzeldaten sichtbar gemacht werden sollen. Jeder Ordner muss durch einen Metadatensatz beschrieben werden, ebenso wie die Dissertationsschrift, die bisher schon separat abgeliefert wird.

Hierfür wurde von der DNB ein Metadaten-Kernset für Forschungsdaten definiert, in dem die obligatorischen und fakultativen Metadaten definiert sind. Es lehnt sich an das DNB-Kernset für Hochschulschriften an und wurde um forschungsdatenspezifische Elemente wie zum Beispiel Relationstypen, Datentypen oder Lizenzinformationen ergänzt.⁶ Ebenso wurde das Kernset für Hochschulschriften erweitert, um Relationen zu Daten von Seiten der Dissertationen abbilden zu können. Die Verweisung erfolgt somit bidirektional.

Die Kernsets dienen zudem als Vorlage für die Erweiterung von XMetaDissPlus, dem DNB-eigenen Metadatenschema für Hochschulschriften. Die Ablieferer sind verpflichtet die Publikationen an der OAI-Schnittstelle mit XMetaDissPlus-Metadaten bereitzustellen, folglich müssen die im Kernset

⁶ Vgl. Deutsche Nationalbibliothek, *Lieferung von Metadaten für monografische Netzpublikationen an die Deutsche Nationalbibliothek*, Version 2.0 (Frankfurt, M.: Deutsche Nationalbibliothek, 2014), 6, [urn:nbn:de:101-2014071100](https://nbn-resolving.org/urn:nbn:de:101-2014071100).

definierten Elemente und Attribute dort auch implementiert werden. Besonderes Augenmerk wurde dabei auf die Abbildung der Relation zwischen Dissertation und Forschungsdaten gelegt, was bisher in XMetaDissPlus nicht möglich war. Die Projektpartner orientierten sich dabei am Metadatenschema von DataCite, das genau diese Aufgabe durch den Einsatz von Relationstypen bereits vorbildlich gelöst hatte.⁷ Diese Relationstypen wurden nunmehr in XMetaDissPlus implementiert und zugleich das Element <relatedIdentifier> mit den entsprechenden Attributen definiert. Die Relation zu einer Dissertation sieht in einem Forschungsdatensatz dann in etwa so aus:

```
<ddb:relatedIdentifier ddb:relatedIdentifierType="URN"  
  ddb:relationType="IsPartOf">  
  urn:beispiel-urn-derDissertation  
</ddb:relatedIdentifier>
```

Vice versa wird diese Relation auch in den Metadaten zu einer Dissertation verzeichnet, dann aber mit dem Relationstyp „HasPart“ und unter Angabe der Forschungsdaten-PID. Die Relationstypen „HasPart“ und „IsPartOf“ sind hierbei verbindlich vorgeschrieben. In der Regel soll als <relatedIdentifierType> eine URN angegeben werden, da in der DNB alle ablieferungspflichtigen Objekte eine URN erhalten. Alternativ ist aber auch die Angabe von DOIs, handles oder anderer PID-Formaten möglich.

Die Relation Dissertation - Forschungsdaten ist eine 1-n-Relation, d.h. einem Dissertationsdatensatz können beliebig viele Forschungsdaten zugeordnet werden. Umgekehrt ist nur eine 1-1-Beziehung möglich. Damit trug man der gesetzlichen Vorgabe Rechnung, dass nur Daten, die Teil einer Veröffentlichung sind, gesammelt werden dürfen und somit jedes Forschungsdatenpaket in Abhängigkeit zu einer bestimmten Dissertation stehen muss.

Dissertationsbezogene Forschungsdaten können somit zukünftig über die von vielen Hochschulen für Dissertationen schon genutzte OAI-PMH-Schnittstelle an die DNB abgeliefert werden – unter der Voraussetzung, dass in den Metadaten die Relation zu einer Dissertation eingetragen ist. Die Metadaten werden in das Katalogsystem der DNB importiert, die Forschungsdaten-Supplemente strikt von der Dissertation getrennt weiterverarbeitet und der Langzeitarchivierung zugeführt. Das alles läuft vollautomatisch ab. Die Supplemente sind nach abgeschlossenem Import als eigenständige Einträge im Katalog recherchierbar; zugleich wird über den Relationstyp im Katalogeintrag die Abhängigkeit zu einer Dissertation nachgewiesen und die Supplement-Funktion der Daten für diese wissenschaftliche Arbeit sichtbar gemacht.

7 DataCite ist ein internationales Konsortium, das sich unter anderem für die einheitliche Registrierung von Forschungsdaten einsetzt und dabei DOI als Objekt-Identifikator verwendet. Für die Registrierung der Daten wurde von DataCite zudem ein Metadatenschema entwickelt. Vgl. DataCite Metadata Working Group., *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data*, Version 4.0 (2016), 18–20, <http://doi.org/10.5438/0012>.

4. Referenzierte Daten

Über dieses Modell lassen sich auch referenzierte Daten ohne Probleme nachweisen. Hierzu wird in den Metadaten der Dissertation eine PID als <relatedIdentifier> eingetragen, die Daten auf externen Repositorien referenziert. Die Daten können nun über diese PID aus dem DNB-Katalog heraus aufgerufen werden. Da ein Dissertationsdatensatz n-Relationen zu Forschungsdaten enthalten darf, können auf diesem Wege beliebig viele Forschungsdaten zu einer Dissertation nachgewiesen werden – ganz gleich ob diese als Supplemente bei der DNB oder auf externen Repositorien vorliegen.

Auf Supplemente wendet die DNB ihren kompletten LZA-Workflow an, d.h. sie führt ein Preservation Planning durch und archiviert die Daten in ihrem Langzeitarchiv.⁸ Bei referenzierten Daten ist jedoch zunächst das jeweilige externe Repository für diese Aufgabe zuständig. Da einige Repositorien sich auf die von der DFG vorgeschlagene 10-jährige Aufbewahrungsfrist⁹ für Forschungsdaten beschränken und keine Langzeitarchivierung durchführen,¹⁰ entwickelte die DNB im Rahmen von eDissPlus ein Konzept, das vorsieht, Daten auf externen Repositorien zu harvesten und sie in den eigenen Langzeitarchivierungs-Workflow einzuspeisen.

5. Zusammenfassung

Das Datenmodell von eDissPlus ermöglicht es der DNB, dissertationsbezogene Forschungsdaten in Zukunft als eigenständige Einträge in ihrem Katalog vorzuhalten und sie damit recherchierbar zu machen.¹¹ Zugleich wird über die Metadaten die Relation zur betreffenden Dissertation abgebildet. Benutzer und Benutzerinnen der DNB können damit zukünftig im Rahmen der gesetzlichen Regelungen im Lesesaal oder online auch Forschungsdaten einsehen. Über die Metadaten der Dissertation wird ihnen auch der Zugriff auf externe Daten ermöglicht, soweit deren PIDs bei Abgabe der Dissertation von den Promovierenden angegeben wurden. Alle Daten, die an die DNB abgeliefert wurden, werden nach Möglichkeit langzeitarchiviert.¹²

Damit leistet die DNB einen Beitrag zur aktuellen Diskussion zu/über Forschungsdaten, die einen Höhepunkt in den 2016 veröffentlichten Empfehlungen des Rates für Informationsinfrastruktur (RFII) fand.¹³ Entsprechend ihrer gesetzlichen Vorgaben sammelt die DNB allerdings nicht alle

8 Vgl. Deutsche Nationalbibliothek, *Langzeitarchivierungs-Policy der Deutschen Nationalbibliothek*, Version 1.0 (Frankfurt, M.: Deutsche Nationalbibliothek, 2013), [urn:nbn:de:101-2013021901](https://nbn-resolving.org/urn:nbn:de:101-2013021901).

9 Vgl. Deutsche Forschungsgemeinschaft, *Vorschläge zur Sicherung guter wissenschaftlicher Praxis: Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft“, ergänzte Auflage* (Weinheim: Wiley-VHC., 2013), 21, <https://doi.org/10.1002/9783527679188.oth1>.

10 Lt. DNB Policy werden derzeit das Data Seal of Approval (DSA) und das nestor-Siegel als vertrauenswürdige Zertifizierungen für Repositorien angesehen. Vgl. Weisbrod, *Policy der Deutschen Nationalbibliothek*, 4.

11 eDissPlus beschäftigt sich nur mit Netzpublikationen. Printpublikationen mit Forschungsdaten wurden nicht bearbeitet.

12 „Je nach Wissenschaftsdisziplin kann es sich bei Forschungsdaten um Messdaten, Beobachtungsdaten, Umfrageergebnisse, oder andere Arten von Daten handeln. Aufgrund der Vielfalt und Komplexität der Datenformate kann die Langzeitarchivierung im Sinne einer unbeschränkten Nutzbarkeit über lange Zeiträume nicht für alle Forschungsdaten von der DNB gewährleistet werden.“ Vgl. Weisbrod, *Policy der Deutschen Nationalbibliothek*, 5.

13 Vgl. Rat für Informationsinfrastrukturen, *Leistung aus Vielfalt: Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland* (Göttingen, 2017), [urn:nbn:de:101:1-201606229098](https://nbn-resolving.org/urn:nbn:de:101:1-201606229098).

Forschungsdaten, sondern nur Supplemente und referenzierte Daten, die ablieferungspflichtigen Netzpublikationen zugeordnet werden können. Damit trägt sie im Sinne der guten wissenschaftlichen Praxis zur Nachprüfbarkeit der in den abgelieferten wissenschaftlichen Arbeiten niedergelegten Forschungsergebnisse bei und ermöglicht deren Nachnutzung. Außerdem wird die DNB die weitere Entwicklung beobachten und bei Bedarf die Lösung von eDissPlus weiterentwickeln. Auf die Möglichkeit, Daten auf Repositorien, die keine langfristige Archivierung anstreben zu identifizieren und zu harvesten, wurde weiter oben schon eingegangen. Zudem wird genau zu beobachten sein, wie die Empfehlungen des Rates für Informationsinfrastrukturen umgesetzt werden und welche Rolle die DNB in einer noch zu schaffenden nationalen Infrastruktur für Forschungsdaten spielen könnte.

Literaturverzeichnis

- Bundesministerium für Justiz und für Verbraucherschutz (BMJV). „Verordnung über die Pflichtablieferung von Medienwerken an die Deutsche Nationalbibliothek“. Zuletzt geprüft am 01.01.2018. <http://www.gesetze-im-internet.de/pflav/index.html>.
- DataCite Metadata Working Group. *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data*. Version 4.0. 2016. <http://doi.org/10.5438/0012>.
- Deutsche Forschungsgemeinschaft. *Vorschläge zur Sicherung guter wissenschaftlicher Praxis: Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft“*. Ergänzte Auflage. Weinheim: Wiley-VHC., 2013. <https://doi.org/10.1002/9783527679188.oth1>.
- Deutsche Nationalbibliothek. *Langzeitarchivierungs-Policy der Deutschen Nationalbibliothek*. Version 1.0. Frankfurt, M.: Deutsche Nationalbibliothek, 2013. <urn:nbn:de:101-2013021901>.
- Deutsche Nationalbibliothek. *Lieferung von Metadaten für monografische Netzpublikationen an die Deutsche Nationalbibliothek*. Version 2.0. Frankfurt, M.: Deutsche Nationalbibliothek, 2014. <urn:nbn:de:101-2014071100>.
- Rat für Informationsinfrastrukturen. *Leistung aus Vielfalt: Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland*. Göttingen, 2017. <urn:nbn:de:101:1-201606229098>.
- Reilly, Susan, Wouter Schallier, Sabine Schrimpf, Eefke Smit und Max Wilkinson. *Opportunities of Data Exchange: Report on Integration of Data and Publications*. 2011. Zuletzt geprüft am 01.01.2018. <http://hdl.handle.net/10013/epic.40198>.
- Weisbrod, Dirk. *Policy der Deutschen Nationalbibliothek für dissertationsbezogene Forschungsdaten*. Version 1.1. Frankfurt, M.: Deutsche Nationalbibliothek, 2017. <urn:nbn:de:101-2017092701>.

Anwendung von OPUS für Forschungsdaten – das Beispiel eines außeruniversitären Forschungsinstituts

Wolfgang Peters-Kottig, Kooperativer Bibliotheksverbund Berlin-Brandenburg

Zusammenfassung:

OPUS wurde als Dokumentenserver konzipiert und ursprünglich nicht für den Nachweis von Forschungsdaten angelegt. Mit vergleichsweise geringem Aufwand können dennoch Forschungsdaten als Supplementary Material zu Zeitschriftenbeiträgen publiziert werden. Auch reine Forschungsdatenpublikationen können via OPUS veröffentlicht werden, selbst wenn die gegenwärtig auf Textpublikationen zugeschnittene Metadatenerfassung weniger gut für die Beschreibung reiner Forschungsdaten geeignet ist. Ein Unterschied zu anderen Systemen und Anwendungsszenarien besteht in der Beschränkung, großvolumige Daten nicht innerhalb von OPUS selbst vorzuhalten. Der für das institutionelle Repository des Zuse-Instituts Berlin konzipierte pragmatische Lösungsansatz ist in OPUS ohne zusätzliche Programmierung umsetzbar und skaliert für beliebig große Forschungsdatensätze.

Summary:

Primarily, OPUS has been designed as a document server and originally was not intended to serve as a repository for research data. However, it is possible to publish research data as supplementary material to a journal article with comparatively simple means. Independent data sets may also be published via OPUS, although the metadata description in OPUS, which has been customized for text publications, is less suited for the description of research data. What makes it different from other systems, however, is the limitation of storage capacity. Large data volumes should not be stored in OPUS. The pragmatic approach designed for the institutional repository of the Zuse Institute Berlin can be implemented in OPUS without additional programming, and is well suited for data sets of any size.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S79-91>

Autorendentifikation: Peters-Kottig, Wolfgang: GND 128736712; ORCID: <http://orcid.org/0000-0003-4486-2422>

Schlagwörter: Repositorien; Forschungsdaten; Publikationsprozess

1. Einleitung

In Deutschland ist OPUS derzeit die am weitesten verbreitete Software für den Betrieb von Repositorien an Hochschulen und Kultureinrichtungen, die eigene Publikationen im Open Access in einem instituts- oder fachbezogenen Repository zur Verfügung stellen wollen.¹ OPUS 4 ist eine vergleichsweise einfach zu administrierende Software, die auch von vielen kleinen Einrichtungen

¹ Laut Directory of Open Access Repositories (zuletzt geprüft am 19.06.2018, www.openoar.org/) werden mit Stand Februar 2018 von 203 gelisteten Repositories in Deutschland 81 mit OPUS betrieben. Die Statistik der KOBV-Zentrale zählt rund 100 OPUS-Installationen in Deutschland (zuletzt geprüft am 19.06.2018, <http://www.kobv.de/entwicklung/software/opus-4/referenzen/>).

mit beschränkteren Ressourcen eingesetzt wird.² Das institutionelle Repository des Konrad-Zuse-Zentrum für Informationstechnik Berlin, kurz Zuse-Institut Berlin oder ZIB, wird ebenfalls mit der Software OPUS betrieben. Seitdem das Thema Forschungsdatenmanagement auch im außeruniversitären Bereich immer größere Bedeutung erlangt und die Relevanz von Forschungsdatenpublikationen zunehmend wahrgenommen wird, gibt es am ZIB Planungen, den Forschenden die Möglichkeit zur Veröffentlichung der eigenen Daten zu geben. Von der Geschäftsleitung des ZIB wurde der Beschluss gefasst, die Möglichkeiten einer Forschungsdatenpublikation in OPUS zu evaluieren und Umsetzungswege zu suchen. Dazu wurde eine gemeinsame Arbeitsgruppe von KOBV (Kooperativer Bibliotheksverbund Berlin-Brandenburg), der ZIB-Abteilung IT & Data Services und interessierten Wissenschaftler/innen des ZIB gegründet. Da die Weiterentwicklung von OPUS 4 seit 2010 maßgeblich von der Verbundzentrale des KOBV am ZIB verantwortet wird, bot sich an, die Planung und eine testweise Implementierung von Funktionalitäten zusammen mit dem OPUS-Team in der haus-eigenen ZIB-OPUS-Instanz durchzuführen.³ Erste Ergebnisse der Arbeiten und ein Ausblick auf das Thema digitale Langzeitarchivierung werden im Folgenden dargestellt.

2. OPUS als Dokumentenserver des ZIB

Das ZIB ist ein außeruniversitäres Forschungsinstitut für angewandte Mathematik und Informatik in Trägerschaft des Landes Berlin. Die Forschungsergebnisse der vielfältigen Kooperationsprojekte werden im ZIB-Repository vor allem in Form von Preprints und Postprints der wissenschaftlichen Fachbeiträge auf Open-Access-Basis veröffentlicht. Daneben gehören Reports und technische Manuskripte sowie Abschlussarbeiten der kooperierenden Berliner Universitäten zu den Inhalten des Dokumentenservers (Abb. 1).

Software-Basis für das Repository ist das Release OPUS 4.6 mit einzelnen ZIB-spezifischen Anpassungen in der Administration und Benutzeroberfläche. So werden nicht nur die OPUS-„Standardfelder“ mit Metadaten zur inhaltlichen Erschließung befüllt.⁴ Zusätzlich werden ergänzende Informationen gesammelt, die eine Einordnung von Dokumenten nach weiteren Kriterien erlauben, beispielsweise danach, ob die betreffende Publikation ein Peer-Review-Verfahren durchlaufen hat. Im Publikationsprozess erfolgt ein automatischer Export der Metadaten in die jeweilige persönliche Publikationsliste auf der ZIB-Webseite.⁵

Der administrative Betrieb von OPUS und die Beratung der Nutzer/innen zählen wie in vergleichbaren Einrichtungen zu den Aufgaben der Institutsbibliothek, wobei das Hosting der Anwendung durch die

2 OPUS 4 ist eine Open-Source-Software unter der GNU General Public License, zuletzt geprüft am 19.06.2018, <https://opus4.kobv.de>. Siehe auch das OPUS-Handbuch unter, zuletzt geprüft am 19.06.2018, <http://www.opus-repository.org/userdoc/>.

3 Zuletzt geprüft am 19.06.2018, <https://opus4.kobv.de/opus4-zib/>

4 Liste der 50 Standard-Metadatenfelder für Dokumente in OPUS: zuletzt geprüft am 19.06.2018, <http://www.opus-repository.org/userdoc/reference/fields.html>. Diese Metadatenfelder basieren auf der DINI-Empfehlung von 2010: DINI Arbeitsgruppe Elektronisches Publizieren, Deutsche Nationalbibliothek und Bibliotheksservice-Zentrum Baden-Württemberg, *Gemeinsames Vokabular für Publikations- und Dokumenttypen*, 2010, <http://doi.org/10.18452/1492>.

5 Beispiel einer automatisch aus den OPUS-Einträgen gespeisten Publikationsliste im ZIB (unter Publications): zuletzt geprüft am 19.06.2018, <http://www.zib.de/members/baum>.

KOBV-Zentrale übernommen wird. Hierbei kann die technische Infrastruktur des ZIB genutzt werden.⁶ Neue Dokumente können nach Login in OPUS in Eigenregie von den Wissenschaftler/inne/n hochgeladen werden. Vor der Freigabe zur Veröffentlichung erfolgt eine Sichtprüfung der eingegebenen Metadaten durch die Institutsbibliothek und es werden Hinweise zur Lizenzvergabe gegeben, z.B. die Empfehlung, Creative-Commons-Lizenzen (bevorzugt CC-BY oder CC0) einzusetzen.



Dokumenttyp	
ZIB-Report	(1529)
Artikel	(1318)
Konferenzbeitrag	(932)
Masterarbeit	(220)
Buchkapitel	(150)
Dissertation	(119)
Sonstiges	(80)
Buch	
(Monographie)	(60)
Beitrag	
Sammelband	(57)
Bachelorarbeit	(48)
ZIB-Jahresbericht	(32)
Poster	(32)
Habilitation	(15)
Bericht	(7)
Konferenzband	(6)
Forschungsdaten	(1)

Abb. 1: Dokumenttypen der Publikationen in der OPUS-Instanz des ZIB. Die Zahlen geben die Anzahl der zugehörigen Dokumente an (Stand Februar 2018).

3. Forschungsdatenmanagement am ZIB

Am ZIB werden große Mengen von Forschungsdaten mit unterschiedlicher Komplexität in einem großen Fächerspektrum mit wissenschaftlichen und industriellen Kooperationspartnern erzeugt. Schwerpunkte liegen in den Lebens-, Natur- und Ingenieurwissenschaften mit der Modellierung, Simulation und Optimierung etwa von Verkehrs- und Versorgungsnetzen und in der Digitalen Medizin. Neben den Herausforderungen der reinen Datenhaltung wird hier die Etablierung eines expliziten Forschungsdatenmanagements (FDM) immer bedeutender.⁷ Bei der Identifizierung der konkreten Herausforderungen für das ZIB ist der Organisationsstatus als verhältnismäßig kleine

⁶ Das ZIB beherbergt einen Supercomputer mit entsprechend leistungsfähiger Hardware-Ausstattung.

⁷ Einen guten Überblick zum aktuellen Stand des Forschungsdatenmanagements in Deutschland geben die Konferenzbeiträge zu den E-Science-Tagen 2017: Jonas Kratzke und Vincent Heuveline, Hrsg., *E-Science-Tage 2017: Forschungsdaten managen* (Heidelberg: heiBOOKS, 2017), <http://doi.org/10.11588/heibooks.285.377>

außeruniversitäre Forschungseinrichtung wichtig, die zudem keiner Wissenschaftsorganisation angehört.⁸ Die Abteilungen innerhalb des ZIB kooperieren über sogenannte Brückenprojekte auch disziplinübergreifend. In einem dieser Brückenprojekte haben unter dem Titel RDM@ZIB vier Wissenschaftler/innen aus drei Abteilungen in Eigeninitiative eine Ad-hoc-Arbeitsgruppe gegründet, die Vorschläge für ein institutionelles FDM entwickelt und umsetzt.⁹

Als eine wichtige Herausforderung und zentrale Chance in Bezug auf ein institutionelles FDM wurde im Rahmen zweier interner Workshops mit Arbeitsgruppenleitern die Datenpublikation identifiziert. Dabei wurde deutlich, dass aufgrund zu geringer Anreize zwar weiterhin wenig Eigeninitiative der Forscherinnen und Forscher im Haus existiert, Daten zu publizieren. Gleichzeitig wäre aber eine Mehrzahl durchaus bereit, eigene Daten zu veröffentlichen unter der Maßgabe, dass Unterstützung im Institut erfolgt und möglichst wenig zusätzliche Arbeitszeit investiert werden muss.

Eine systematische Befragung aller AGs im ZIB ist noch nicht erfolgt, die Ergebnisse der Workshops und einzelner, unstrukturierter Interviews mit Arbeitsgruppenleitern geben aber einen Eindruck der bestehenden Probleme. Derzeit besteht eine Datenveröffentlichung am ZIB in Aufzeichnungen auf persönlichen oder Projekt-Homepages, die keine recherchierbare, standardisierte Metadatenbeschreibung, keine Versionierung und keine nachhaltige Bereitstellung erlauben. Für den Datenaustausch mit Dritten werden häufig auf Arbeitsgruppenebene Server betrieben und es werden (versteckte) Webseiten erzeugt, auf denen die Daten von den Partnern wieder abgeholt werden können. Hier fehlt es – abgesehen von hardware- und softwarebedingten Anforderungen – an Unterstützung bei der Metadatenbeschreibung. Dies betrifft sowohl deskriptive Metadaten im Fachkontext als auch technische und administrative Metadaten. Authentifizierungs-/Autorisierungsmechanismen und eindeutige Identifizierung von Objekten in allen Stationen des Datenlebenszyklus werden ebenfalls noch nicht standardmäßig eingesetzt. Die bisherige ZIB-Praxis bezogen auf nachhaltige Verfügbarkeit von Projektergebnissen als Forschungsdaten besteht überwiegend darin, nach Ende eines Forschungsprojekts oder dem Ausscheiden verantwortlicher Personen (insbesondere Doktorand/inne/en), unpublizierte Daten auf einen gesonderten Archivbereich des Dateisystems auszulagern. Die Übertragbarkeit dieser Bestandsaufnahme auf die Archivierungsbedarfe von originären Forschungsprojekten am ZIB ist aus technischer, organisatorischer und finanzieller Perspektive noch ungeklärt.

Im Ergebnis hat die AG RDM@ZIB zwei vorrangige Herausforderungen aus dem Status quo identifiziert:

- Beseitigung der Defizite bei Datenpublikationsmöglichkeiten, Erhöhung der Publikations-Anreize und verbesserte Information über vorhandene Optionen.
- Standardisierung von Metadatenbeschreibungen und Schaffung geregelter Workflows zu Datentransfer und -analyse, Datenablage, Archivierung sowie Datenweitergabe.

8 Am ZIB sind rund 230 Personen beschäftigt.

9 Eine Initialzündung für den Start des institutionellen FDM war die ZIB-Beteiligung am DFG-geförderten Projekt EWIG; s. die im Projekt entstandene Broschüre: Roland Bertelmann et al., *Einstieg ins Forschungsdatenmanagement in den Geowissenschaften* (Potsdam, 2014), <http://doi.org/10.2312/lis.14.01>.

Die zweite Herausforderung ist deutlich komplexer und wird in diesem Kontext (mit Ausnahme der Archivierung) nicht näher beleuchtet, weil sie erst in einem späteren Schritt angegangen werden soll. Fraglich ist z.B., wo die Grenze zwischen Aufgaben des institutionellen FDM und Hardwarebereitstellung als Aufgabe der „IT-Infrastruktur“ bzw. dem Forschungsgegenstand des ZIB liegt.¹⁰

Die Datenpublikation mit OPUS wurde als wichtigste Aufgabe zuerst angegangen. Ein wichtiger begrenzender Faktor ist hier die Organisationsstruktur des ZIB als kleine Forschungseinrichtung mit geringen Personalkapazitäten für Service-Aufgaben. Deshalb war es wichtig, an einer praktikablen Lösung zu arbeiten, die in der ersten Annäherung möglichst ohne zusätzliche Entwicklungsarbeiten OPUS „datengerecht“ anpassen kann und nur geringe Änderungen im Workflow vorsieht. Noch nicht geklärt ist für die ZIB-Instanz die Frage, ob und in welchem Umfang eine weitergehende, vielleicht sogar fachliche Qualitätsprüfung vor Veröffentlichung erfolgen kann.¹¹ Ein Upload soll auch bei Forschungsdaten möglichst eigenständig durch die Beitragenden selbst erfolgen können, wie es bereits bei Textpublikationen in der derzeitigen OPUS-Instanz der Fall ist. Parallel wurde ein initialer Beratungsservice durch die AG aufgebaut, um die Interessentinnen und Interessenten bei allen Fragen zum Thema Data Publishing zu unterstützen – konkret auch bei der Erarbeitung von Datenmanagementplänen.

4. Publikation von Forschungsdaten mit OPUS

Der vergleichsweise geringe administrative Aufwand wird im ZIB als großer Vorteil von OPUS gesehen. Aus den langjährigen Erfahrungen der KOBV-Verbundzentrale – nicht zuletzt im Austausch mit externen Repository-Betreibern – wurde in der AG der Schluss gezogen, dass komplexe Software-Alternativen wegen eines zu erwartenden hohen Konfigurations- und Wartungsaufwands nicht in Frage kommen würden. Gleichzeitig sollte ein Betrieb des Repositoriums weder physikalisch noch „ideell“ ausgelagert werden, so dass etwa Zenodo oder figshare als potentielle Publikationsorte ebenfalls ausschieden. Als Vorteil beim Einsatz von OPUS 4 kann gelten, dass die Forscherinnen und Forscher bereits bekannte Workflows nutzen können. Zudem kann lokal vorhandene Infrastruktur eingesetzt werden. Publikationen und Forschungsdaten können an einer zentralen Stelle im ZIB verknüpft und bereitgestellt werden. OPUS 4 besitzt zudem ein flexibles, auch lokal erweiterbares Metadatenschema.

Ideen und Umsetzungsvorschläge für die Verzeichnung von Forschungsdaten mit OPUS bestehen schon seit längerem, etwa bei den Entwicklungspartnern von OPUS, der UB Stuttgart und dem Bibliotheksservice-Zentrum Baden-Württemberg.¹² Basierend auf diesen Vorarbeiten wurden Gespräche

10 Die Bereitstellung einer Plattform für hochperformante Analyse und Verarbeitung von großvolumigen Daten in Echtzeit wird von der AG RDM@ZIB nicht zum Aufgabenbereich des institutionellen FDM gezählt. Die Metadatenbeschreibung in diesem Kontext aber sehr wohl – dafür gibt es jedoch noch kein tragfähiges Konzept am ZIB.

11 Wie von Van Tuyl und Whitmire (2016) empfohlen, sollten „Repository Manager“ zumindest prüfen können, ob hinreichend fachspezifische Metadaten vorhanden und die Forschungsdaten grundsätzlich nachnutzbar sind. Vgl. Steven Van Tuyl und Amanda L. Whitmire, „Water, Water, Everywhere: Defining and Assessing Data Sharing in Academia,” *PLoS ONE* 11, Nr. 2 (2016), e0147942: 14, <https://doi.org/10.1371/journal.pone.0147942>.

12 Matthias Schulze, „Virtuelle Forschungsumgebungen und Forschungsdaten für Lehre und Forschung: Informationsinfrastrukturen für die (Natur-)Wissenschaften,” in *Semantic Web & Linked Data : Elemente zukünftiger*

mit dem OPUS-Team beim KOBV geführt, um technische Grundlagen und funktionale Anforderungen zu klären und eine Testumsetzung zu implementieren. Die folgende Definition der Funktionalität von Datenpublikationen diene als Grundlage für die Implementierung in der ZIB-OPUS-Instanz:

a) Publikation als Supplementary Material. In OPUS wird bei einer Textpublikation auf zusätzlich vorhandenes Supplementary Material verwiesen. Dazu steht ein neues Metadatenfeld „URL der Forschungsdaten“ zur Verfügung sowie ein neues Metadatenfeld „Bemerkungen zu Forschungsdaten“ mit der Möglichkeit, die verlinkte Ressource zu beschreiben. Das Feld „URL der Forschungsdaten“ verweist entweder zu einer externen Quelle, wie einem Forschungsdaten-Repository/einer Verlagsseite/einem Cloudspeicher, oder zu einem eigenständigen Forschungsdaten-Eintrag in OPUS mit eigener Landingpage, wenn die Forschungsdaten auf einem ZIB-Server vorgehalten werden (und über einen persistenten Identifier, z.B. DOI verfügen). Auf dieser zusätzlichen Landingpage ist ein Enrichment-Feld „Data Download“ vorhanden, das bei Anklicken den eigentlichen Download ermöglicht.

b) Eigenständige Forschungsdatenpublikation. Wenn keine zugehörige Textpublikation in OPUS vorhanden ist, kann auch ein eigenständiges Dokument „Forschungsdaten“ angelegt werden. Die Forschungsdaten werden entweder außerhalb des ZIB bei einer beliebigen Quelle oder wie unter a) genannt auf einem eigenen ZIB-Server vorgehalten. Im Feld „URL der Forschungsdaten“ wird dann entweder auf die externe Ressource verlinkt oder unter „Data Download“ das Herunterladen vom ZIB-Server ermöglicht. Eine Beschreibung der Ressource ist im Feld „Bemerkungen zu Forschungsdaten“ möglich. Mit dem Feld „URL“ ist auch ein Rückverweis auf z.B. später erschienene Textpublikationen zum Forschungsdatensatz möglich.

Der erste Arbeitsschritt bestand in der Festlegung notwendiger Erweiterungen zum Metadatenchema von OPUS. Anhand von Supplementary Material zu einer aktuellen Publikation zweier ZIB-Mitarbeiter wurden auf Basis des DataCite-Metadatenchemas und im Abgleich mit weiteren Datenschemata wie dem RADAR Metadata Schema die Felder identifiziert, für die kein Mapping von vorhandenen Standardfeldern in OPUS sinnvoll oder möglich ist.¹³ Für diese Felder können in OPUS vergleichsweise einfach (ohne Programmieraufwand) sogenannte Enrichment-Felder in der Administrationsschicht angelegt werden. Diese Möglichkeit gilt für alle OPUS 4-Betreiber. Abb. 2 listet für die ZIB-Instanz vorläufig neu definierte Enrichment-Felder auf. Hier gilt es abzuwägen, welche Felder wirklich generisch für alle Fachdisziplinen gelten und welche eher fachspezifisch sind. Weil OPUS beim Anlegen von Enrichments sehr flexibel ist, wurde vorläufig entschieden, bei Bedarf

Informationsinfrastrukturen; Proceedings der 1. DGI-Konferenz, 62. Jahrestagung der DGI, Frankfurt am Main, 7.-9. Oktober 2010, Tagungen der DGI 14, hrsg. Marlies Ockenfeld (Frankfurt am Main, 2010), 172, <http://dx.doi.org/10.18419/opus-6344>.

Einen aktuellen Umsetzungsvorschlag des BSZ zur Forschungsdatenpublikation in OPUS hat Friederike Gerland auf dem OPUS-Workshop beim Bibliothekskongress 2017 präsentiert: Friederike Gerland, „Forschungsdatenmanagement in OPUS 4“ (Vortrag auf dem 106. Deutschen Bibliothekartag in Frankfurt am Main 2017), Vortragsfolien, <http://nbn-resolving.de/urn:nbn:de:0290-opus4-30859>.

- 13 DataCite Metadata Working Group, *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data*. Version 4.1 (2017), <http://doi.org/10.5438/0014>.
RADAR Projektteam, *Dokumentation des deskriptiven RADAR Metadatenchemas*. Version 0.9 (FIZ Karlsruhe, 2017), zuletzt geprüft am 14.05.2018, https://www.radar-service.eu/sites/default/files/RADAR_Metadaten_Dokumentation_v09.pdf

auch sehr spezifische Erweiterungen zuzulassen.¹⁴ Gleichzeitig wurde ein neuer Dokumenttyp „Forschungsdaten“ (Research Data) in der OPUS-Administration angelegt, der das Browsen in der Instanz bezogen auf Forschungsdaten ermöglicht (s. auch Abb. 1).

Enrichmentfeld	Benennung en.	Benennung dt.
EnrichmentScientificNote	Notes to research data	Bemerkungen zu den Forschungsdaten
EnrichmentScientificName	Scientific name	Wissenschaftlicher Name
EnrichmentScientificDateCollected	Date data collected	Datum der Datenerhebung
EnrichmentScientificDateCreated	Date data created	Datum der Datenerstellung
EnrichmentScientificGeolocation	Geolocation	Ort der Datenerhebung
EnrichmentScientificResourceTypeGeneral	ResourceTypeGeneral	ResourceTypeGeneral
EnrichmentRelatedIdentifier	Related identifier	Related Identifier
EnrichmentScientificDataDownload	Research data download	URL zum DataDownload
EnrichmentSoftwareDescription	Software description	Software Description

Abb. 2: Neu definierte Metadatenfelder (Enrichments) für Forschungsdaten in OPUS.

Bei der Verwendung von OPUS für Forschungsdatenpublikationen gibt es in der aktuellen Version folgende Einschränkungen bzw. Herausforderungen.

- Aufgrund der Architektur von OPUS ist es nicht sinnvoll, große Datenvolumina innerhalb von OPUS zu verwalten. Der Upload mehrerer 100 GB – wie bei Forschungsdaten im ZIB durchaus gängig – ist nur auf Umwegen möglich weil die Administration größtenteils über eine Web-Oberfläche erfolgt.
- Eine Auslagerung der eigentlichen Download-Daten auf einen externen Server (außerhalb der OPUS-Verwaltung) ist sinnvoll, um Performance-Probleme mit großvolumigen Forschungsdaten zu vermeiden. In diesem Fall kann aber das Rollen- und Rechtesystem von OPUS nicht genutzt werden. Diese Funktionalität sollte von einem eigenständigen Authentifizierungssystem abgedeckt werden.
- Eine Vorschau ist nicht möglich. Forschungsdaten können nicht direkt im OPUS-Repository geöffnet oder gestreamt werden. Stattdessen muss immer die gesamte Datei lokal heruntergeladen werden.
- Für bestimmte Datentypen, insbesondere Software, ist es wichtig, dass auf alle Versionen verwiesen werden kann (Vorläuferversion, Nachfolgeversion, Landingpage zum Haupteintrag für die Software). OPUS bietet aktuell keine detaillierte Versionierung für publizierte Objekte.

¹⁴ Das Feld „Scientific name“ ist beispielsweise sehr spezifisch. Es bezeichnet in dem zugrundeliegenden Datensatz aus der Bioinformatik den taxonomischen Namen einer Tierart.

- Wissenschaftler/innen wollen zu Recht die arbeitsintensive, manuelle Mehrfacheingabe von Metadaten wie Autor/Principal Investigator, Titel, Förderkennzeichen u.ä., in Forschungsinformationssystemen, Repositories, Projektmanagementsoftware etc. vermeiden. Dieser Mehraufwand lässt sich derzeit mangels Schnittstellen zwischen heterogenen Systemen kaum vermeiden.

Die genannten Herausforderungen wurden als unkritisch für den gewünschten Einsatz von OPUS erachtet oder können umgangen werden. Die Performance-Einschränkung bei der Verwaltung großer Datenmengen wird vermieden, wenn Forschungsdatenobjekte (Dateien) grundsätzlich außerhalb von OPUS in einer eigenen Infrastruktur verwaltet werden. Auf der Landingpage (Frontdoor) in OPUS selbst wird nur ein Download-Link hinterlegt, der auf einen unabhängigen Downloadserver innerhalb des ZIB verweist (Abb. 3). Hilfreich ist hier die spezifische Organisation des ZIB, welches ein eigenes Rechenzentrum inklusive Hardware-Infrastruktur in Form der Abteilung IT & Data Services beherbergt. Somit stellte sich vor Ort nicht das Problem der Vertrauenswürdigkeit und Nachhaltigkeit eines extern betriebenen Downloadservers.¹⁵

Alternativ kann an dieser Schnittstelle auch ein separates Rollen- und Rechtemanagement vorgeschaltet werden, das im einfachsten Fall eine Passwortabfrage aufruft oder an ein Verfahren zur verteilten Authentifizierung und Autorisierung wie Shibboleth angebunden werden kann.¹⁶ Eine Einschränkung des Zugangs im Open Access gilt es allerdings sorgfältig abzuwägen. Das ZIB verfolgt eine Open-Access-Strategie, die sich etwa in der langjährigen Praxis niederschlägt, (Zweit-)Veröffentlichungen von ZIB-Angestellten als sogenannte ZIB-Reports im Open Access herauszugeben.¹⁷ Es wird angestrebt, dass auch Forschungsdaten möglichst nicht mit einem Embargo oder anderen Restriktionen versehen werden.¹⁸ Gleichzeitig ist anzumerken, dass ein dem Download vorgeschaltetes Rechtemanagement auch zur Vereinfachung eines Peer-Review-Verfahrens beitragen kann. Gutachtern einer Datenpublikation lässt sich im Rahmen des Review-Prozesses auf diesem Weg flexibel der Zugriff auf Daten im Repository freischalten.

Die oben genannte eingeschränkte Möglichkeit zur Versionsangabe wird in Kauf genommen weil Aktualisierungen mit Verweisen auf Vorgänger- und Nachfolgeversionen durch die OPUS-Administration manuell in Enrichment-Feldern eingetragen werden können. Dies ist zwar umständlich, aber jedenfalls hinreichend aussagekräftig (eine Lösung für Versionenverwaltung befindet sich in der Entwicklungsplanung für zukünftige OPUS-Releases). In der Metadatenanzeige auf der Landingpage zu dem Ende 2017 veröffentlichten Testdatensatz sind einige noch nicht umgesetzte kleinere Probleme sichtbar, die v.a. der von den Wissenschaftlern erwünschten kurzfristigen Veröffentlichung geschuldet

15 Für OPUS-Betreiber, die nicht auf entsprechende institutionelle Infrastrukturen zurückgreifen können, kommen ggf. Cloudspeicher-Services als Lösung in Frage.

16 Maßgeblich ist hier der weitverbreitete Dienst des Deutschen Forschungsnetzes DFN-AAI: zuletzt geprüft am 19.06.2018, <https://www.aai.dfn.de/>.

17 Auch die im Oktober 2017 verabschiedete Open-Access-Leitlinie des KOBV am ZIB spiegelt diesen Anspruch wider: Kooperativer Bibliotheksverbund Berlin-Brandenburg, *Open Access als strategische Leitlinie des KOBV*, ZIB-Report 17-54, 2017. <http://nbn-resolving.de/urn:nbn:de:0297-zib-65303>.

18 Aus forschungsethischen und/oder datenschutzrechtlichen Gründen kann eine Zugriffsbeschränkung nicht immer vermieden werden, etwa im Fall der am ZIB etablierten Verarbeitung medizinischer Forschungsdaten.


waren (Abb. 3 bzw. der Originaldatensatz unter <https://doi.org/10.12752/4.DKN.1.0>). Vorteilhaft ist wiederum die derzeit in Entwicklung befindliche automatische DOI-Vergabe in OPUS.¹⁹ Bei der Vorbereitung der Veröffentlichung des Testdatensatzes mussten die Meldung an DataCite und das Eintragen des DOI noch manuell erfolgen.


Der Link auf die zugehörige Veröffentlichung ist noch nicht direkt anklickbar und rechts oben auf der Landingpage ist noch kein Download-Icon angezeigt, sondern der Dateidownload erfolgt über den Link im Feld „Research data download“, was etwas unübersichtlich erscheint.


Die grundlegende Aufgabe, eine Funktionalität in OPUS umzusetzen, mit der eine zitierfähige, persistente Forschungsdatenpublikation mit geringem personellem Aufwand von geschätzt weniger als einem Personenmonat möglich ist, war aus Sicht des ZIB aber sehr erfolgreich. Mehrere ZIB-Arbeitsgruppen haben Anfang 2018 bereits Interesse an einer kurzfristigen Datenpublikation geäußert. Der Ansatz soll nach dem Prinzip „Learning by Doing“ laufend optimiert und ausgebaut werden.²⁰ Für Einrichtungen, die OPUS 4 ebenfalls für ihr institutionelles Repository einsetzen, dürfte der Lösungsweg des ZIB für die Publikation von Forschungsdaten mit vergleichbar geringem Aufwand umsetzbar sein.

¹⁹ Die Integration dieser Funktionalität ist für eines der kommenden OPUS 4-Releases noch in 2018 vorgesehen.

²⁰ Beispielsweise ergaben sich wertvolle Hinweise für das OPUS-Team für die sich derzeit in Entwicklung befindliche automatisierte DOI-Vergabe. Beim Mapping der OPUS-Metadaten auf das DataCite MD-Schema sollte etwa der Feldinhalt von „GeoLocation“, der bei Forschungsdaten den Ort der Original-Datenerhebung bezeichnet, nicht mit dem Verlagsort belegt werden.

 Automated Segmentation of Complex Patterns in Biological Tissues: Lessons from Stingray Tessellated Cartilage (Supplementary Material)



 [David Knötel](#), [Ronald Seidel](#), [Paul Zaslansky](#), [Steffen Prohaska](#), [Mason N. Dean](#), [Daniel Baum](#)

 Supplementary data to reproduce and understand key results from the related publication, including original image data and processed data. In particular, sections from hyomandibulae harvested from specimens of round stingray *Urobatis halleri*, donated from another study (DOI: 10.1002/etc.2564). Specimens were from sub-adults/adults collected by beach seine from collection sites in San Diego and Seal Beach, California, USA. The hyomandibulae were mounted in clay, sealed in ethanol-humidified plastic tubes and scanned with a Skyscan 1172 desktop μ CT scanner (Bruker μ CT, Kontich, Belgium) in association with another study (DOI: 10.1111/joa.12508). Scans for all samples were performed with voxel sizes of 4.89 μ m at 59 kV source voltage and 167 μ A source current, over 360° sample 120 rotation. For our segmentations, the datasets were resampled to a voxel size of 9.78 μ m to reduce the size of the images and speed up processing. In addition, the processed data that was generated with the visualization software Amira with techniques described in the related publication based on the mentioned specimens.


EXPORT METADATA

[BibTeX](#) [RIS](#) [XML](#)

ADDITIONAL SERVICES

[Send a mail to the author of this document](#)

 **Metadaten**


Author:	David Knötel , Ronald Seidel , Paul Zaslansky , Steffen Prohaska , Mason N. Dean , Daniel Baum
Document Type:	Research data
ResourceTypeGeneral:	Dataset
Year of first publication:	2017
Notes to research data :	Supplementary data to reproduce and understand key results from the related publication, including original image data and processed data.
Date data created:	2017
Geolocation:	San Diego and Seal Beach, California, USA
Scientific name :	<i>Urobatis halleri</i>
Related identifier:	https://doi.org/10.1371/journal.pone.0188018
Software description :	The data was processed with the visualization software Amira. See http://www.zib.de/software/tesseract-segmentation for information about Amira and how to download the Amira extension package created for this publication. See https://github.com/zibamira/tesseract-segmentation.git for the source code of the Amira extension package.
Research data download:	http://www.zib.de/ext-data/2017_Knoetel_PLOSONE-Supplementary-Material.zip
DOI:	https://doi.org/10.12752/4.DKN.1.0
Licence (German):	 Creative Commons - Namensnennung

Abb. 3: Landingpage (in OPUS „Frontdoor“) zum ersten als Supplementary Material veröffentlichten Forschungsdatensatz im ZIB.

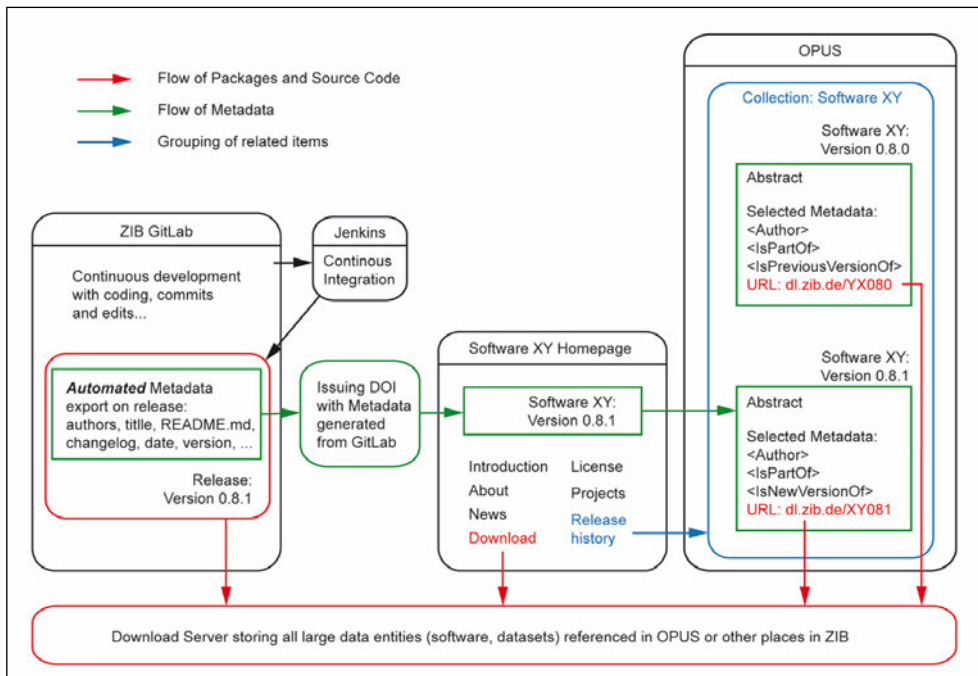


Abb. 4: Architekturkonzept bei Veröffentlichung von Software in OPUS.

5. Ausblick: Softwarepublikation und Langzeitarchivierung

Weil im ZIB mathematische Software eine wichtige Ressource und ein Forschungsprodukt darstellt und zukünftig als Forschungsdatum auch in OPUS publizierbar sein soll, gibt es erste Ideen für die Verknüpfung der Software-Entwicklung mit dem Veröffentlichungsworkflow (Abb. 4). Dabei soll ebenfalls der oben skizzierte Ansatz verfolgt werden, große Datenpakete nicht innerhalb von OPUS selbst zu verwalten, sondern diese auf eine getrennte Infrastruktur (innerhalb des ZIB) auszulagern. Der adressierte Server muss aber sicherstellen, dass keine nachträglichen Änderungen einmal publizierter Daten möglich sind, sondern Versionen separat publiziert werden. Kernaufgaben werden nach diesem Konzept die automatische Metadatenübernahme aus der Entwicklungsumgebung (GitLab) nach OPUS und die Abbildung der verschiedenen Versionen einer Software in OPUS sein. Dabei soll auch das Problem der mangelnden Zuschreibung einer Werkautorschaft für Software bearbeitet werden. In den ZIB-AGs, die Software entwickeln, herrscht große Personalfuktuation und die Projekte dauern häufig sehr lange. Dies gestaltet die angemessene Attribuierung komplizierter als bei herkömmlichen Text- und Forschungsdatenpublikationen.²¹ Zudem wird Software in der Mathematik und den Ingenieur- und Naturwissenschaften bislang häufig gar nicht oder nicht

21 Laura Soito und Lorraine J. Hwang, „Citations for Software: Providing Identification, Access and Recognition for Research Software.“ *International Journal of Digital Curation* 11, Nr. 2 (2016): 53, <http://doi.org/10.2218/ijdc.v11i2.390>.

separat zitiert und die Zitationen sind diffus, d.h. Software wird auf unterschiedliche Weise genannt (Publikationen, Manuals, Webseiten, etc.). Für dieses disziplinübergreifende Problem der heterogenen Zitationspraxis von Software sollen in Kooperation mit swMATH²² Lösungsvorschläge für die Mathematik erarbeitet werden.

Die im Datenlebenszyklus definierte Aufgabe der Archivierung hat im Sinn einer dauerhaften Nachnutzbarkeit eine der Publikation vergleichbare Bedeutung. In der ZIB-Instanz von OPUS soll die Archivierung als explizite, OAIS-konforme digitale Langzeitarchivierung an das Langzeitarchiv EWIG im ZIB ausgelagert werden.²³ Ein automatisierter Export aus OPUS nach EWIG erfolgt idealerweise in einem METS-Container; dabei kann der bereits bestehende Exportpfad mittels XMetaDisPlus als Basis für den Metadaten austausch dienen. In Abhängigkeit von der Entwicklungsroadmap des OPUS-Teams soll diese Funktionalität zukünftig in Kooperation mit der ZIB-AG „Digital Preservation“ entwickelt werden.

Für andere Repositorien kann es ebenfalls empfehlenswert sein, die eigentliche Forschungsdatenarchivierung unabhängig von der Online-Präsentation zu betreiben, weil dann kein komplexes „Alleskönner-System“ mit nur einer Software implementiert und gewartet werden muss. In OPUS selbst müssen durch die klare Trennung von Publikation und Langzeitarchivierung keine komplexen Qualitätssicherungswerkzeuge für die Datenprüfung²⁴ in den Veröffentlichungsworkflow implementiert werden und die Objekt- bzw. Metadatenverwaltung von OPUS lässt sich vollständig auf das eigentliche Veröffentlichen fokussieren.

Literaturverzeichnis

- Bertelmann, Roland, Petra Gebauer, Tim Hasler, Ingo Kirchner, Wolfgang Peters-Kotting, Matthias Razum, Astrid Recker, Damian Ulbricht und Stephan van Gassel. *Einstieg ins Forschungsdatenmanagement in den Geowissenschaften*. Potsdam, 2014. <http://doi.org/10.2312/lis.14.01>.
- DataCite Metadata Working Group. *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data*. Version 4.1, 2017. <http://doi.org/10.5438/0014>.

22 swMATH ist ein in Zusammenarbeit von ZIB, Forschungscampus MODAL und FIZ Karlsruhe betriebenes Informationssystem für mathematische Software, zuletzt geprüft am 19.06.2018, <https://www.swmath.org/>.

23 Der KOBV und das ebenfalls am ZIB angesiedelte Forschungs- und Kompetenzzentrum Digitalisierung Berlin (digiS) entwickeln und betreiben gemeinsam das digitale Langzeitarchiv EWIG; zuletzt geprüft am 19.06.2018, <http://www.kobv.de/services/archivierung/lza/>. Für technische Details zur Architektur von EWIG siehe: Marco Klindt und Kilian Amrhein, „One Core Preservation System for All your Data. No Exceptions!“, iPRES 2015 - Proceedings of the 12th International Conference on Preservation of Digital Objects, 101–108 (School of Information and Library Science, University of North Carolina at Chapel Hill, 2015), <http://hdl.handle.net/11353/10.429551>.

24 Solche Werkzeuge zur Dateiformatidentifizierung und -validierung sind Teil des Ingest-Prozesses des Langzeitarchivs.

- DINI, Arbeitsgruppe Elektronisches Publizieren, Deutsche Nationalbibliothek und Bibliothekservice-Zentrum Baden-Württemberg. *Gemeinsames Vokabular für Publikations- und Dokumenttypen*. 2010. <http://doi.org/10.18452/1492>.
- Gerland, Friederike. „Forschungsdatenmanagement in OPUS 4“. Vortrag auf dem 106. Deutschen Bibliothekartag in Frankfurt am Main 2017. Vortragsfolien. <http://nbn-resolving.de/urn:nbn:de:0290-opus4-30859>.
- Klindt, Marco und Kilian Amrhein. „One Core Preservation System for All your Data. No Exceptions!“, iPRES 2015 - Proceedings of the 12th International Conference on Preservation of Digital Objects, 101–108. School of Information and Library Science, University of North Carolina at Chapel Hill, 2015. <http://hdl.handle.net/11353/10.429551>.
- Kooperativer Bibliotheksverbund Berlin-Brandenburg. *Open Access als strategische Leitlinie des KOBV*. ZIB-Report 17–54, 2017. <http://nbn-resolving.de/urn:nbn:de:0297-zib-65303>.
- Kratzke, Jonas und Vincent Heuveline, Hrsg. *E-Science-Tage 2017: Forschungsdaten managen*. Heidelberg: heiBOOKS, 2017. <http://doi.org/10.11588/heibooks.285.377>.
- RADAR Projektteam. *Dokumentation des deskriptiven RADAR Metadatenschemas*. Version 0.9. FIZ Karlsruhe, 2017. Zuletzt geprüft am 14.05.2018. https://www.radar-service.eu/sites/default/files/RADAR_Metadaten_Dokumentation_v09.pdf.
- Schulze, Matthias. „Virtuelle Forschungsumgebungen und Forschungsdaten für Lehre und Forschung: Informationsinfrastrukturen für die (Natur-)Wissenschaften“. In *Semantic Web & Linked Data : Elemente zukünftiger Informationsinfrastrukturen*; Proceedings der 1. DGI-Konferenz, 62. Jahrestagung der DGI, Frankfurt am Main, 7.–9. Oktober 2010. Tagungen der DGI 14, herausgegeben von Marlies Ockenfeld, 165–176. Frankfurt am Main, 2010. <http://dx.doi.org/10.18419/opus-6344>.
- Soito, Laura und Lorraine J. Hwang. „Citations for Software: Providing Identification, Access and Recognition for Research Software.“ *International Journal of Digital Curation* 11, Nr. 2 (2016): 48–63. <http://doi.org/10.2218/ijdc.v11i2.390>.
- Van Tuyl, Steven und Amanda L. Whitmire. „Water, Water, Everywhere: Defining and Assessing Data Sharing in Academia.“ *PLoS ONE* 11, Nr. 2 (2016), e0147942. <https://doi.org/10.1371/journal.pone.0147942>.

SARA-Dienst: Software langfristig verfügbar machen

Franziska Rapp, Stefan Kombrink, Volodymyr Kushnarenko, Universität Ulm

Matthias Fratz, Daniel Scharon, Universität Konstanz

Zusammenfassung:

Software spielt in vielen Disziplinen eine wichtige Rolle im Forschungsprozess. Sie ist entweder selbst Gegenstand der Forschung oder wird als Hilfsmittel zur Erfassung, Verarbeitung und Analyse von Forschungsdaten eingesetzt. Zur Nachvollziehbarkeit der durchgeführten Forschung sollte Software langfristig verfügbar gemacht werden. Im SARA-Projekt zwischen der Universität Konstanz und der Universität Ulm wird ein Dienst entwickelt, der versucht die Einschränkungen bereits bestehender Angebote aufzuheben. Dies beinhaltet u.a. die Möglichkeit, die gesamte Entwicklungshistorie auf einfache Weise mitzuveröffentlichen und für Dritte zur Online-Exploration anzubieten. Zudem bestimmen die Forschenden den Zeitpunkt und Umfang der zu archivierenden/veröffentlichenden Software-Artefakte selbst. Der SARA-Dienst sieht auch die Möglichkeit vor, eine Archivierung ohne Veröffentlichung vorzunehmen. Der geplante Dienst verbindet bereits bestehende Publikations- und Forschungsinfrastrukturen miteinander. Er ermöglicht aus der Arbeitsumgebung der Forschenden heraus eine Archivierung und Veröffentlichung von Software und unterstützt Forschende dabei, bereits prozessbegleitend Zwischenstände ihrer Forschung festzuhalten. Aufgrund seines modularen Aufbaus kann der SARA-Dienst in unterschiedlichen Szenarien zum Einsatz kommen, beispielsweise als kooperativer Dienst für mehrere Einrichtungen. Er stellt eine sinnvolle Ergänzung zu bestehenden Angeboten im Forschungsdatenmanagement dar.

Summary:

Software plays an important role in many scientific disciplines. Whether software itself is the research focus or whether software tools are used to create, process and analyse data – software should be available for the long term to make the research process reproducible. In the context of the SARA project conducted by the University of Konstanz and Ulm University, a service is being developed which aims to avoid restrictions of existing services. This includes the possibility to easily publish the whole change history and make it available for others to explore online. Additionally, the researchers decide when and what they want to archive/publish. The SARA service also allows for archiving of software without making it publicly accessible. It connects existing publication and research infrastructures. Researchers can trigger a publication of software from their research environment and are encouraged to publish software artefacts already during the research process. The new service can be used in various scenarios due to its modular design, for example as a cooperative service for several institutions. The SARA Service is a useful addition to already existing research data management services.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S92-105>

Autorenidentifikation: Rapp, Franziska: GND 1131101308; Fratz, Matthias: GND 1097829235; Kombrink, Stefan: GND 1150699604; Scharon, Daniel: GND 1093687045; Kushnarenko, Volodymyr: GND 1147751196

Schlagwörter: Softwarearchivierung, Veröffentlichung von Software, SARA-Service, DSpace, GitLab, Institutionelles Repositorium, Reproduzierbarkeit, Git

1. Motivation

Software spielt seit Jahrzehnten in wissenschaftlicher Forschung in vielen Disziplinen eine essentielle Rolle. Ob zur Erfassung, Analyse und Interpretation von Forschungsdaten oder als eigenständiges Forschungsdatum – Software sollte zur Nachvollziehbarkeit der durchgeführten Forschung längerfristig archiviert und verfügbar gemacht werden.¹ Die Spanne an Software reicht dabei von kleinen Skripten oder Makros über umfangreiche und komplexe Programme bis hin zu kompletten Software-Frameworks. Wir sprechen in diesem Kontext auch von Software-Artefakten.

Um sich im Rahmen von Publikationen auf Software-Artefakte beziehen zu können, sollten diese entsprechend zitierbar sein. Die Software Citation Principles der FORCE 11 „Software Citation Working Group“ stellen unter anderem die folgenden Anforderungen an Software-Publikationen:²

- Sie sollen durch einen Persistent Identifier zitierbar sein (am besten DOI).
- Diese sollen auf eine jeweilige Software-Version verweisen.
- Eine Landing Page mit zusätzlichen Informationen vor dem eigentlichen Download sollte angezeigt werden.

Um all diesen Anforderungen gerecht zu werden, möchte das Projekt SARA (Software Archiving of Research Artefacts) einen Dienst aufbauen, über den im Rahmen von Forschung entwickelte und angepasste Software-Artefakte nach diesen Vorgaben publiziert werden können. Dabei werden bereits vorhandene Entwicklungs- und Publikationsinfrastrukturen nachgenutzt und über SARA miteinander verbunden.

Der Dienst soll die Forschenden in ihren Workflows begleiten und sie anregen, Zwischenstände ihrer Forschungsarbeit auch bei Software-Artefakten bereits prozessbegleitend festzuhalten. Zur besseren Nachvollziehbarkeit der Forschung sollte die Entwicklungshistorie miterhalten werden. Dies ist insbesondere von Bedeutung, wenn man die übliche Vorgehensweise bei der Softwareentwicklung berücksichtigt, nach welcher komplexere Software-Artefakte nicht komplett eigenständig neuentwickelt werden, sondern auf bestehenden Entwicklungen aufbauen. So verwendet jedes nicht-triviale Programm Programmbibliotheken oder erweitert/ergänzt bestehenden Programmcode.

2. Stand der Technik

In der Software-Entwicklung gilt es seit langem als Best Practice, sämtlichen Quellcode in einem Versionsverwaltungssystem zu hinterlegen, in dem jede Änderung protokolliert und dokumentiert wird. Hierüber werden Änderungen verschiedener Entwickler/innen zusammengeführt und auftretende Konflikte gelöst. Durch diesen Prozess entsteht eine lückenlose Dokumentation des Entstehungsprozesses eines Software-Artefakts, anhand derer sich jede Zwischenversion wiederherstellen

1 Arbeitskreis Open Science, „Helmholtz Open Science: Zugang zu und Nachnutzung von wissenschaftlicher Software.“ 2017, zuletzt geprüft am 12.02.2018, <https://os.helmholtz.de/?id=2766>.

2 Arfon M. Smith et al., „Software Citation Principles,“ *PeerJ Computer Science* 2 (2016): e86, 12-13, <https://doi.org/10.7717/peerj-cs.86>.

lässt. Im wissenschaftlichen Bereich ermöglicht dies die Reproduzierbarkeit softwaregestützter Datenverarbeitung.

Git bildet unter Versionsverwaltungssystemen mittlerweile den De-facto-Standard, insbesondere bei neuen Projekten, gefolgt von Subversion (SVN), das noch in vielen älteren Projekten genutzt wird. Von diesen beiden Systemen ist Git besser zur Archivierung geeignet: Es ist eine neuere Konzeption mit aktiver Entwicklung und wird damit voraussichtlich länger verfügbar sein. Insbesondere ist Git von vornherein als verteiltes System ausgelegt, mit explizit systemunabhängigen, genau definierten Klartextformaten. SVN-Repositories können zudem leicht in Git-Repositories umgewandelt werden. Jede Arbeitskopie eines Git-Repositories enthält eine vollständige Kopie der gesamten Versionsgeschichte und ist somit, anders als SVN, nicht von einem Server oder Dienst abhängig, auf dem die kanonische Kopie der Versionsgeschichte liegt.

Für Open-Source-Projekte gibt es zahlreiche Anbieter öffentlicher Git-Repositories. Die bekanntesten sind GitHub³, SourceForge⁴ und Google Code⁵ (eingestellt). Die Bereitstellung durch größere Unternehmen bietet eine bessere Persistenz als bei einem lokalen Git-Repository – beispielsweise hat Google bei der Einstellung von Google Code sämtliche Projekte explizit archiviert. Diese Dienste sind aber nicht von vornherein als Archiv ausgestaltet und garantieren in der Regel keine dauerhafte Speicherung, d.h. Projekte könnten grundsätzlich das Schicksal von GeoCities teilen, einem Dienst, der 2009 eingestellt wurde, ohne dass von Anbieterseite eine Archivierung erfolgte.⁶

Mit Zenodo⁷ gibt es einen Dienst zur Archivierung von Software, der vom CERN gehostet wird und mit dessen Hilfe beliebige Versionsstände eines Projekts auf GitHub zitierbar (mittels DOI) archiviert und veröffentlicht werden können. Hierüber kann jederzeit der aktuelle Stand der Software zitierbar archiviert werden. Dadurch ist die Reproduzierbarkeit der mit dieser Version verarbeiteten Forschungsdaten gewährleistet. Es wird jedoch nur die jeweilige Version selbst gespeichert. Die Versionsgeschichte kann lediglich im ursprünglichen Projekt auf GitHub nachverfolgt werden. Dieses ist auf Zenodo verlinkt, wird jedoch nicht mitarchiviert – das Projekt kann vom Nutzer nach wie vor jederzeit verändert oder (auch versehentlich) gelöscht werden.

Darüber hinaus existieren zahlreiche weitere, teils fachspezifische Repositorien für Forschungsdaten und Software. Der Schwerpunkt liegt zumeist nicht auf Software. Source Code kann als Ordner oder als ZIP-Archiv desselben hinterlegt werden. Wenn archivierende Nutzer/innen darauf achten, den versteckten git-Ordner mit einzuschließen, bleibt dabei die Versionsgeschichte enthalten und kann nach dem Entpacken genutzt werden. Das Ergebnis, ein ZIP-Archiv einer Working Copy, hat aber den Charakter einer Notlösung. Eine Web-Voransicht, sofern vorhanden, zeigt in der Regel lediglich Dateien und keine Versionen an. Metainformationen wie Commit Messages werden nicht angezeigt.

3 GitHub, <https://github.com>.

4 SourceForge, <https://sourceforge.net/>.

5 Google Code Archive, <https://code.google.com/archive/>.

6 Christian Klauß, „Geocities geschlossen - und für die Nachwelt archiviert.“ *Golem.de*, 26. Oktober 2009, zuletzt geprüft am 12.02.2018, <https://www.golem.de/0910/70706.html>.

7 Zenodo, <https://zenodo.org/>.

Zenodo kommt den gewünschten Eigenschaften am nächsten: Für die archivierten Versionsstände ist eine zuverlässige Speicherung durch das CERN gewährleistet. Zudem kann bei noch aktiven Projekten die Versionsgeschichte im zugehörigen GitHub-Projekt exploriert werden. Der Erhalt dieser Daten ist jedoch nicht garantiert. Insofern fehlt eine Kombination zur garantierten, mittel- bis langfristigen Speicherung der gesamten Versionsgeschichte, kombiniert mit der Möglichkeit diese bequem betrachten und analysieren zu können.

Das Projekt SARA – Software Archiving of Research Artefacts wurde Anfang 2016 als Kooperation der Arbeitsgruppe Verteilte Systeme an der Universität Konstanz, dem Institut für Organisation und Management von Informationssystemen an der Universität Ulm und dem Kommunikations- und Informationszentrum der Universität Ulm gestartet und verfolgt das Ziel, einen neuen Dienst zur Archivierung und Publikation von Software anzubieten, der diese Einschränkungen bestmöglich aufhebt.

3. Der SARA-Workflow: Archivieren & Publizieren von Software

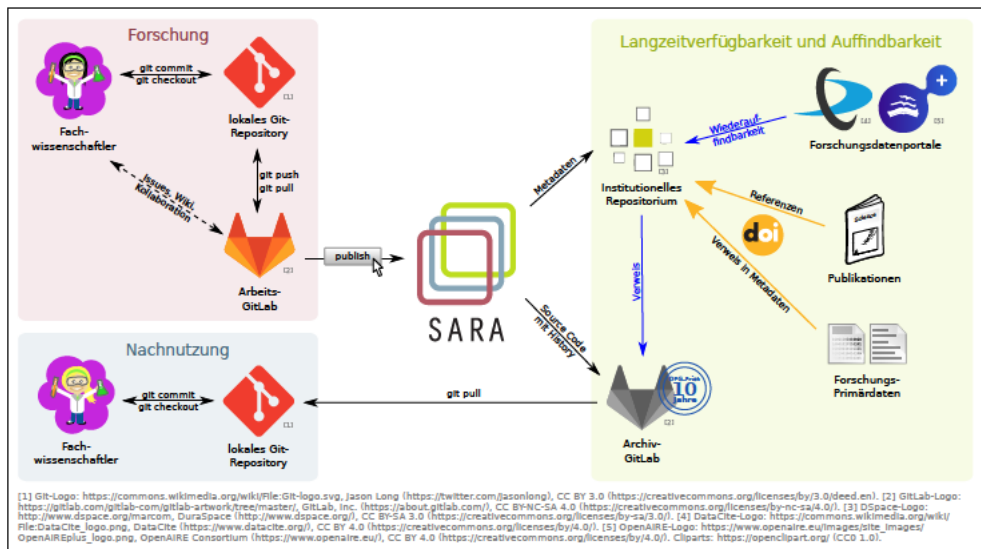


Abb. 1:⁸ Die verschiedenen Komponenten des SARA-Service. Der SARA-Dienst ist so konzipiert, dass Arbeits-GitLab, Archiv-GitLab und institutionelle Repositorien mehrfach vorkommen können.

8 Franziska Ackermann et al., „SARA-Service: Langzeitverfügbarkeit und Publikation von Softwareartefakten“ (Poster auf den E-Science-Tagen in Heidelberg März 2017), <https://doi.org/10.11588/heidok.00022887>, CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>.

Der SARA-Workflow in Kürze

Der SARA-Server ist der Dreh- und Angelpunkt für die Archivierung und Veröffentlichung von Software. Er verbindet die verschiedenen Komponenten miteinander. Zu Beginn steht ein Git-Repository auf z.B. einer GitLab-Instanz, in dem die Forschenden einzeln oder kollaborativ an Projekten arbeiten und direkt aus dieser Umgebung heraus eine Archivierung und Veröffentlichung anstoßen können. Wir bezeichnen diese Arbeitsumgebung an dieser Stelle als „Arbeits-GitLab“. Der SARA-Server nimmt den Auftrag der Forschenden entgegen, bietet ihnen verschiedene Archivierungs- und Veröffentlichungsoptionen an und fragt weitere Details ab. Anschließend wird die Überführung der Software-Artefakte in ein Archiv-GitLab angestoßen, welches die Langzeitverfügbarkeit sicherstellt. Die URL, unter der die Daten zu finden sind, wird zusammen mit weiteren Metadaten in einem institutionellen⁹ Repositorium veröffentlicht. Über die Repositorien erfolgt die Vergabe einer DOI oder eines anderen Persistent Identifier zur dauerhaften und eindeutigen Adressierung der veröffentlichten Software-Artefakte. Der SARA-Dienst berücksichtigt, dass nicht alle Software-Artefakte publiziert werden können (z.B. aus urheberrechtlichen, ethischen, patentrechtlichen Gründen) und bietet daher auch die Möglichkeit, eine reine Archivierung vorzunehmen, bei der die Daten nicht öffentlich zugänglich sind. Über die Repositorien sind Metadaten zu den Software-Artefakten auch in verschiedenen Nachweissystemen wie OpenAIRE¹⁰, BASE¹¹, Google Scholar¹², DataCite Search¹³ etc. zu finden. Dies erhöht die Wiederauffindbarkeit der Daten und ermöglicht es anderen Forschenden, Software-Artefakte bei entsprechender Lizenzierung aus dem Archiv-GitLab in die eigene Forschungsumgebung zu übernehmen und nachzunutzen.

Der große Mehrwert von SARA für Forschende liegt darin, dass sie aus ihrer gewohnten Forschungsumgebung heraus eine Archivierung und optional Veröffentlichung vornehmen können und dabei viele Schritte automatisiert ablaufen. Gleichzeitig liegen relevante Entscheidungen weiterhin bei den Forschenden.

Merkmale des SARA-Service

- Forschende bestimmen Zeitpunkt und Umfang der Veröffentlichung selbst (im Unterschied zu automatisierten Veröffentlichungen ohne Selektionsmöglichkeit)
- Archivierung der Entwicklungshistorie möglich (im Unterschied zu einer Verlinkung auf dynamische, löschbare Projekte)
- Sicherstellung der Langzeitverfügbarkeit über Archiv-GitLab
- Online-Datenexploration im Archiv-GitLab
- Sicherstellung der Auffindbarkeit und Zitierbarkeit über DOI-Vergabe in institutionellen Repositorien

9 Der geplante Dienst kann ebenso zusammen mit fachlichen Repositorien eingesetzt werden.

10 OpenAIRE, <https://www.openaire.eu/>.

11 BASE Bielefeld Academic Search Engine, <https://www.base-search.net/>.

12 Google Scholar, <https://scholar.google.de/>.

13 DataCite, <https://search.datacite.org/>.

Der SARA-Workflow im Detail

3.1. Forschung

Softwareentwicklung findet zunächst in einem lokalen Git-Repository statt, das sich z.B. auf dem Rechner der Forschenden befindet. Über einen kollaborativ genutzten Git-Server erfolgt die Zusammenarbeit mit anderen Forschenden. Bei lokal an Einrichtungen oder Universitäten betriebenen Diensten kommt in der Regel die Software GitLab zum Einsatz. GitLab bietet ein Web-Frontend, das eine grafische Oberfläche und weitere Funktionalitäten für die Zusammenarbeit bietet (Issues, Wikis etc.). Im Folgenden sprechen wir deshalb von einem „Arbeits-GitLab“. Der SARA-Dienst ist so strukturiert, dass eine Anbindung an verschiedene Git-Web-Frontends möglich ist. Eine zusätzliche Anbindung an GitHub wird im Rahmen des Projekts erprobt.

Die Grundlage im SARA-Workflow bildet das Arbeits-GitLab. Darin können einzelne Forscher/innen in einem eigenen Git-Repository arbeiten, z.B. wenn es sich um eine eigenständig durchgeführte Abschlussarbeit handelt, oder es können kollaborativ mehrere Forschende an einem gemeinsamen Projekt arbeiten. Das „Arbeits-GitLab“ unterscheidet sich vom „Archiv-GitLab“ dadurch, dass es für die gemeinsame Zusammenarbeit genutzt wird, während das „Archiv-GitLab“ die Langzeitverfügbarkeit sicherstellt und für Standard-Nutzer nur lesende Rechte beinhaltet.

Für die produktive und effiziente Nutzung eines Versionsverwaltungssystems haben sich in der Softwareentwicklung Best Practices entwickelt. Diese Verhaltensregeln und Workflows sind zunächst besonders für die Softwareentwicklung geeignet, können aber als Grundlage für andere Fachdisziplinen dienen. Sofern einige grundlegende und wenig strittige Konventionen eingehalten werden, kann dadurch auch die Archivierung und Publikation der Software erheblich vereinfacht werden. Dazu zählt insbesondere, Commits mit sinnvollen Metadaten zu versehen – mit dem korrekten Namen der Nutzerin oder des Nutzers oder einem offiziellen Alias, einer offiziellen (bevorzugt institutionellen) E-Mail-Adresse und einer aussagekräftigen Beschreibung der Änderungen. Git unterstützt Nutzer/innen dabei dahingehend, dass ihre Identität gespeichert wird und Commit-Beschreibungen (in sehr eingeschränktem Umfang) generiert werden können.

3.2. SARA-Server

Auf dem SARA-Server können Forschende das GitLab auswählen, in dem sie arbeiten, und ein Projekt selektieren, das ganz oder in Teilen archiviert/veröffentlicht werden soll. Anschließend werden sie auf dem SARA-Server durch verschiedene Schritte geleitet, zu denen die Bestimmung des Umfangs und der Art der Archivierung/Veröffentlichung gehören sowie die Auswahl einer Lizenz und Vergabe von Metadaten.

Forschende nehmen eine Selektion der zu archivierenden/veröffentlichenden Daten vor und entscheiden sich für eine bestimmte Archivierungs- oder Veröffentlichungsoption. Dabei kann einerseits nach Branches selektiert werden, so dass Irrwege, verworfene Ideen oder nicht veröffentlichbare Daten aussortiert werden können, andererseits kann die Versionsgeschichte auch auf die wichtigsten Schritte reduziert werden. Diese Selektion erfolgt auf dem SARA-Server und ist von der verwendeten Repository-Software unabhängig.

SARA SOFTWARE ARCHIVING OF RESEARCH ARTEFACTS

How much do you want to publish?

← What will the options do?

- **Publish full history** publishes all commits in the version history, even those excluded in previous publications (if any). It does not change commit IDs, and is the **recommended** option because it allows replicating your research process.
- **Publish abbreviated history** publishes only **significant points** of the version history and thus **changes commit IDs**. Significant points are defined to be merge commits and commits that have been tagged with **release**. This is useful if you need to omit parts of your development process.
- **Publish latest version only** publishes only the **HEAD revision**, omitting the version history except for previously published versions. This option is **not recommended** if you have other options.
- **Private access, public record** privately archives the full history of a branch and publishes a bibliographical record only. This is suitable for data that you want to (or have to) keep private.
- **Private access, no record** privately archives the full history of a branch. The existence of the archived data will not be published.

Please note that none of these options are suitable for highly sensitive data. If you have that kind of data, please archive it elsewhere.

branch master: starting at most recent commit (HEAD of this branch) | publish full history (recommended) | remove

branch webapp: 398c041 and before: update version to 3.1415 (2017-06-14) | publish abbreviated history | remove

tag test: starting at most recent commit (HEAD of this branch) | private access, public record | remove

Any other branches you would like to include?

tag a-b-c | add

ⓘ Merged branches and tags on selected branches are always included when publishing full or abbreviated history. You do not have to add them explicitly.

← save and return to git repo | next ->

Abb. 2: Auswahl von Umfang und Art der Archivierung/Veröffentlichung

Im Bereich der Metadaten werden Informationen automatisch aus Git und GitLab extrahiert, damit Nutzer/innen möglichst wenig selbst eingeben müssen. Orientieren sich Nutzer/innen bei der Arbeit mit Git an bestimmten Best Practices, erfahren sie an dieser Stelle eine Arbeitserleichterung. Eine der Best Practices im Open-Source-Bereich sieht vor, im Wurzelverzeichnis des Git-Repository eine Datei mit dem Namen „LICENSE“ oder „COPYING“ anzulegen, in welcher die Lizenz als Text enthalten ist. Diese Datei wird vom SARA-Dienst automatisch erkannt. Ist keine Datei hinterlegt, müssen Nutzer/innen eine solche Datei anlegen oder während des Archivierungs- /Publikationsprozesses eine Lizenz auswählen, die vom SARA-Dienst als LICENSE-Datei zu den archivierten Daten geschrieben wird. Der SARA-Dienst stellt auf diese Weise sicher, dass eine Lizenz vergeben wurde, die interessierte Dritte als LICENSE oder COPYING-Datei im Wurzelverzeichnis des Git-Repository vorfinden. Somit sind die Möglichkeiten und Grenzen einer Nachnutzung von Anfang an explizit geklärt. Was technisch nicht garantiert werden kann ist, dass Nutzer/innen eine sinnvolle Lizenz vergeben. Dies muss z.B. bei Abschlussarbeiten mit der Betreuerin oder dem Betreuer besprochen werden. Auch Beratungsstellen der Einrichtungen können hier Hilfestellung geben.

Nach erfolgreicher Archivierung der ausgewählten Daten im Archiv-GitLab folgt als nächster Schritt die Veröffentlichung über ein institutionelles Repository, sofern anfangs eine Option ausgewählt wurde, die eine Veröffentlichung beinhaltet. In diesem Fall geben die Nutzer/innen ihre E-Mail-Adresse

an, mit welcher sie im Repository registriert sind. Dadurch kann der Nachweis in ihrem Namen angelegt werden und die spätere Kommunikation zwischen Repositorien und Nutzer/inne/n wird erleichtert. Für den Nachweis der Software-Artefakte übermittelt der SARA-Dienst die erfassten Metadaten über eine geeignete Schnittstelle ins Repository. Da veröffentlichte Software-Artefakte einen DOI erhalten sollen, werden hierfür notwendige Pflichtmetadaten vom SARA-Server als Minimalmetadatensatz abgefragt. Hiervon ausgenommen sind der Identifier, der erst im Repository vergeben wird und die Angabe des Publisher, welche die betreibende Einrichtung selbst vornimmt. Zum Minimalmetadatensatz gehören außerdem der Link ins Archiv-GitLab und die Person, welche die Veröffentlichung angestoßen hat.

DataCite Pflichtfelder: [Identifier]; Creator; Title; [Publisher]; Publication Year; Resource Type

Weitere Pflichtfelder: Link ins Archiv-GitLab; Person, welche die Veröffentlichung angestoßen hat

Im Projekt liegt der Fokus auf der Anbindung von DSpace¹⁴, das weltweit als Software für institutionelle Repositorien eingesetzt wird, Open Source ist und eine große und aktive Community besitzt. Sowohl an der Universität Ulm als auch an der Universität Konstanz wird DSpace als institutionelles Repository eingesetzt und ist Veröffentlichungsplattform und Universitätsbibliographie zugleich. Das Konzept des SARA-Dienstes sieht einen modularen Ansatz vor, sodass theoretisch auch die Anbindung weiterer Repositorien erfolgen kann, die auf einer anderen Software basieren.

Für DSpace sieht der SARA-Dienst vor, dass verschiedene Varianten der Veröffentlichung unterstützt werden, je nach Philosophie der Einrichtung:

Variante 1 (Workspace): Der SARA-Dienst stößt im Namen der Nutzerin oder des Nutzers eine Submission in DSpace an, die zunächst in den Bereich des Nutzers gelangt (sog. „Workspace“). Die Nutzerin oder der Nutzer loggt sich im institutionellen Repository ein, ergänzt die Metadaten ggf. um institutionsspezifische Felder, die ihm in der DSpace-Instanz seiner Einrichtung angeboten werden, stimmt ggf. einem Veröffentlichungsvertrag zu und schließt den Vorgang ab. Sofern in der DSpace-Instanz ein Überprüfungsschritt vor der Freischaltung eingerichtet ist (sog. „Workflow“), ist der Nachweis nicht sofort öffentlich sichtbar, sondern erst nach einer formalen Prüfung durch die Sachbearbeitung. Dieser zusätzliche Schritt ist in DSpace nicht erforderlich, kann jedoch eingerichtet werden. Dies wird von den meisten institutionellen Repositorien mit von Nutzer/inne/n eingereichten Submissions zu Zwecken der Qualitätskontrolle (in der Regel Normierung und Ergänzung von Metadaten) so praktiziert.

Variante 2 (Workflow): Diese Variante setzt voraus, dass in DSpace der Überprüfungsschritt durch die Sachbearbeitung eingerichtet ist. Der SARA-Dienst stößt eine Submission in DSpace an, die direkt in den Bearbeitungspool für die Sachbearbeitung gelangt (sog. „Workflow“). Die Sachbearbeiter/innen prüfen das Item, ergänzen ggf. Metadaten und treten bei Fragen mit der Nutzerin oder dem Nutzer in Kontakt. Eine Ergänzung & Korrektur der Metadaten in DSpace erfolgt durch die

¹⁴ DSpace, <http://www.dspace.org/>.

Sachbearbeiter/innen, da die Nutzerin oder der Nutzer keinen Zugriff auf die Daten hat. Möchte eine Einrichtung, dass Nutzer/innen selbst neben dem Minimalmetadatensatz weitere Felder ausfüllen, so ist vorgesehen, dass die Einrichtung den SARA-Dienst so konfigurieren kann, dass Nutzer/inne/n bereits dort weitere Metadatenfelder ausfüllen können oder müssen. Nach einer formalen Prüfung wird der Nachweis freigeschaltet und ist öffentlich sichtbar.

Variante 3 (Archiv): Wenn weder Variante 1 noch Variante 2 zum Tragen kommen soll, gelangen Nachweise, die vom SARA-Server an DSpace übermittelt werden, direkt in das sog. DSpace-„Archiv“, was in der Regel bedeutet, dass sie direkt öffentlich sichtbar sind. Die Entscheidung hierüber liegt beim Betreiber des institutionellen Repositoriums.

3.3. Langzeitverfügbarkeit und Auffindbarkeit

Die Langzeitverfügbarkeit der archivierten bzw. veröffentlichten Software-Artefakte wird über ein Archiv-GitLab sichergestellt. Dort werden die archivierten Daten für mindestens zehn Jahre vorgehalten, entsprechend den Empfehlungen der DFG zur „Sicherung guter wissenschaftlicher Praxis“.¹⁵ Als Exitstrategie ist derzeit geplant, im Bedarfsfall sämtliche Dateien in einer ausgecheckten Working Copy des Git-Repository als ZIP-Datei auf einem geeigneten Dienst zur Archivierung abzulegen. Die Versionsgeschichte bleibt dadurch erhalten.

Die Auffindbarkeit wird über institutionelle Repositorien sichergestellt, die zur dauerhaften und eindeutigen Adressierung der veröffentlichten Software-Artefakte DOIs oder andere Persistent Identifier vergeben. Sie können in Publikationen oder anderen Veröffentlichungen für das Zitieren der Software-Artefakte verwendet werden. Über den DOI gelangt man zur Landing Page im Repository, auf der beschreibende Metadaten und der Link ins Archiv-GitLab zu finden sind. Über den Link erfolgt der Sprung ins Archiv-GitLab zu den Software-Artefakten. Diese können im Webbrowser exploriert werden und bei Bedarf in die eigene Arbeitsumgebung der Forschenden übernommen werden.

Über die bei Repositorien bereits etablierten Mechanismen zur Verbreitung von Metadaten wird die Wiederauffindbarkeit der veröffentlichten Software-Artefakte erhöht (EU-Portal OpenAIRE, BASE, DataCite Search, Data Citation Index¹⁶, Google Scholar u.a.).

3.4. Nachnutzung

Veröffentlichte Software-Artefakte können im Archiv-GitLab online exploriert werden. Es besteht die Möglichkeit zum Download der archivierten Version sowie der Zwischenversionen (sofern die Entwicklungshistorie mitveröffentlicht wurde). Eine „LICENSE“ oder „COPYING“ Datei im Git-Repository informiert über die Nutzungsbedingungen. Gegebenenfalls können die Software-Artefakte für weiterführende Forschung in die eigene Arbeitsumgebung übernommen werden.

15 „Deutsche Forschungsgemeinschaft, *Sicherung guter wissenschaftlicher Praxis: Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft“*“, Ergänzte Auflage. (Weinheim: WILEY-VCH Verlag GmbH & Co. KGaA, 2013). <http://dx.doi.org/10.1002/9783527679188.oth1>.

16 „Data Citation Index,“ Clarivate Analytics, zuletzt geprüft am 15.02.2018, http://wokinfo.com/products_tools/multidisciplinary/dci/.

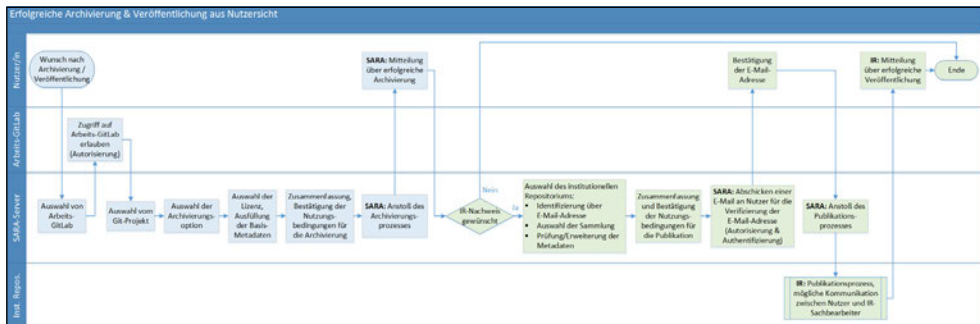


Abb. 3: Der Archivierungs-/Veröffentlichungsprozess aus Nutzersicht

4. Einsatzszenarien für den SARA-Dienst: individuell oder kooperativ

Durch die geplante Bereitstellung der im SARA-Projekt vorgenommenen Entwicklungen als Open Source ergeben sich verschiedene Nachnutzungsmöglichkeiten.

Ein mögliches Szenario für einen SARA-Dienst kann der Einsatz innerhalb nur einer Einrichtung sein: SARA-Server, Arbeits-GitLab, institutionelles Repository und Archiv-GitLab werden von derselben Einrichtung betrieben. Der Vorteil liegt darin, dass alles aus einer Hand angeboten wird, die Wege kurz sind und keine Vereinbarungen mit Dritten erfolgen müssen. Dem steht entgegen, dass die Gefahr besteht, dass über die Zeit Einzellösungen entstehen und Synergien durch kooperative Lösungen ungenutzt bleiben.

Die Stärke des Konzepts für einen SARA-Dienst liegt darin, dass der Dienst kooperativ genutzt werden kann. Das E-Science-Projekt SARA fasst einen solchen Dienst für die Universitäten und Hochschulen im Land Baden-Württemberg ins Auge. Der geplante Dienst ist im Kontext der E-Science-Strategie des Landes Baden-Württemberg zu sehen.¹⁷ Ein Arbeits-GitLab kann als Landesdienst angeboten werden, die Authentifizierung würde über bwIDM¹⁸ erfolgen. Alternativ oder zusätzlich können bestehende Arbeits-GitLabs, z.B. einzelner Institute, an den SARA-Dienst angebunden werden. Interessierte Einrichtungen sollen das eigene institutionelle Repository (ggf. auch mehrere) an den SARA-Dienst anschließen können. Zunächst ist dies nur für DSpace-basierte Repositorien geplant. Durch den modularen Ansatz kann dies jedoch theoretisch auf weitere Repositorien-Software ausgeweitet werden. Forschende teilnehmender baden-württembergischer Universitäten und Hochschulen können im skizzierten Szenario aus dem Arbeits-GitLab heraus archivieren und veröffentlichen. Der Nachweis der archivierten Software-Artefakte und die Vergabe eines Persistent Identifier erfolgen im institutionellen Repository der jeweiligen Einrichtung. Für Forschende der Universität Ulm erfolgen

17 „E-Science: Wissenschaft unter neuen Rahmenbedingungen,“ Ministerium für Wissenschaft, Forschung und Kunst, zuletzt geprüft am 12.02.2018, <https://mwk.baden-wuerttemberg.de/de/forschung/forschungslandschaft/e-science/>.

18 „bwIDM | Föderiertes Identitätsmanagement der baden-württembergischen Hochschulen,“ <https://www.bwidm.de/>.

Nachweis/Veröffentlichung und DOI-Registrierung beispielsweise in OPARU¹⁹, für Forschende der Universität Konstanz in KOPS²⁰. Als Archiv-GitLab soll eine fürs ganze Land bereitgestellte Installation dienen. Der SARA-Server übernimmt die Kommunikation der Komponenten untereinander. Der Vorteil der kooperativen Nutzung liegt darin, dass Arbeits-GitLab, Archiv-GitLab und SARA-Server nicht von jeder Einrichtung einzeln betrieben werden müssten. Der geplante kooperative Dienst orientiert sich an den Infrastruktur-Empfehlungen des Positionspapiers „Zugang zu und Nachnutzung von wissenschaftlicher Software“ des Arbeitskreises Open Science der Helmholtz-Gemeinschaft.²¹

5. Architektur

Die Grafik stellt die Architektur des SARA-Dienstes dar. Da der Dienst als eine Web-Anwendung konzipiert ist, durchläuft ein Nutzer den Archivierungs- und Nachnutzungsvorgang stets mithilfe seines Web-Browsers.

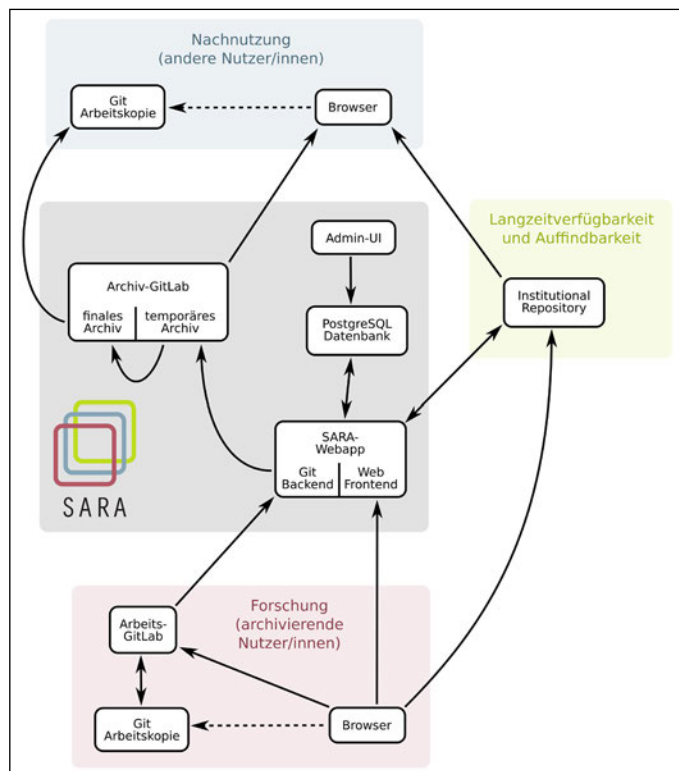


Abb. 4: Die Architektur des SARA-Dienstes

19 OPARU, <https://oparu.uni-ulm.de/>.

20 KOPS – Das Institutionelle Repositorium der Universität Konstanz, <https://kops.uni-konstanz.de/>.

21 Arbeitskreis Open Science, „Helmholtz Open Science: Zugang zu und Nachnutzung von wissenschaftlicher Software.“ Punkt 4.4. Infrastrukturen, 2017, zuletzt geprüft am 12.02.2018, <https://os.helmholtz.de/?id=2766>.

Die Web-Anwendung (SARA WebApp) benötigt zum Betrieb eine Datenbank, die für teilnehmende Einrichtungen zuvor konfiguriert werden muss. Dort lassen sich die Zugangsdaten für die Archiv-GitLabs, institutionellen Repositorien und deren detaillierte Konfiguration wie Default-Kollektionen und Metadaten-Mappings voreinstellen. Zudem werden dort die benutzerdefinierten Metadaten zwischengespeichert, bis der Archivierungs- bzw. Publikationsvorgang abgeschlossen wurde.

Der Web-Dienst wird in Java EE und JavaScript realisiert. Für die Daten wird eine PostgreSQL Datenbank eingesetzt, welche lokal oder remote betrieben werden kann. Vorerst wird die Anbindung an die institutionellen Repositorien mittels SWORD Protokoll und REST Interface umgesetzt, was voraussetzt, dass teilnehmende Einrichtungen für beide Schnittstellen einen Benutzerzugang freischalten müssen.

6. Laufzeitumgebung: Kooperation mit CiTAR

Das SARA-Projekt adressiert die Langzeitverfügbarkeit und Nachweisbarkeit von Software-Artefakten, welche im Rahmen von Forschungsprojekten entwickelt wurden und ermöglicht, dass diese zitiert werden können. Zu Recht ist eine zeitgleiche Archivierung der von den Entwicklern verwendeten Laufzeitumgebung häufig ein Muss-Kriterium, damit die ursprünglichen Ergebnisse der Software verifiziert werden können. Die längerfristige Archivierung von Laufzeitumgebungen wie kompilierte Software, Tool Chains und Betriebssysteminstallationen sprengt die Kapazitäten des SARA-Projektes. Jedoch ist dies eine zentrale Aufgabe des Projektes „Citing and archiving research“ (CiTAR), welches ebenso wie SARA im E-Science-Bereich angesiedelt ist. Für eine mögliche Zusammenarbeit wurden die folgenden Punkte identifiziert:

Forschende können im Git-Projekt eine „Bauanleitung“ für die Entwicklungsumgebung mitführen und -pflegen (Best Practice). Findet SARA beim Publikations- oder Archivierungsvorgang eine solche Anleitung, ist angedacht, dass die Nutzung des CiTAR-Dienstes zur Erstellung und längerfristigen Archivierung der Entwicklungsumgebung angeboten wird. Statt einer Bauanleitung können auch „natürlich gewachsene“ Arbeitsumgebungen an den CiTAR-Dienst übergeben werden. Forschende müssen sicherstellen, dass ihre Entwicklungsumgebung „self-contained“ ist, d.h. keine externen Ressourcen-Abhängigkeiten aufweist. CiTAR gibt das Datenformat vor, in dem die Umgebungen vorliegen müssen, um importiert werden zu können. Derzeit werden die Containerlösungen Docker und Singularity unterstützt, eine Unterstützung für Virtuelle-Maschinen-Abbilder ist geplant. CiTAR unterstützt nur Linux-basierte Entwicklungsumgebungen.

Die Kooperation beider Projekte zielt darauf ab, dass archivierte Software-Artefakte auch zu einem späteren Zeitpunkt wieder in ihrer ursprünglichen Laufzeitumgebung nachnutzbar sind. Die Chance auf eine 1:1-Reproduzierbarkeit wird dadurch erheblich erhöht.

7. Fazit

Mit der Entwicklung des SARA-Dienstes sollen Software-Artefakte, die im Rahmen von Forschung entstehen, bereits prozessbegleitend archiviert und publiziert werden können. Dabei wird darauf geachtet, dass u.a. die eingangs genannten Anforderungen und Empfehlungen von FORCE 11 und des Arbeitskreises Open Science der Helmholtz-Gemeinschaft berücksichtigt werden. Wichtige Merkmale des SARA-Dienstes sind, dass Zeitpunkt und Umfang der Archivierung/Veröffentlichung selbst bestimmt werden können. Dabei ist auch eine Archivierung ohne Veröffentlichung möglich. Die Änderungshistorie kann zudem ebenfalls mitarchiviert werden und auf einfache Weise mitpubliziert und für Dritte online zur Exploration angeboten werden. Das modulare Konzept des geplanten SARA-Dienstes ermöglicht verschiedene Einsatzszenarien, beispielsweise als kooperativ betriebener Landesdienst und stellt eine sinnvolle Ergänzung zu bestehenden Services einer Einrichtung im Bereich des Forschungsdatenmanagements dar.

Acknowledgements

Das SARA-Projekt dankt dem Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg für die Förderung des Projektes.

Literaturverzeichnis

- Ackermann, Franziska, Petra Enderle, Matthias Fratz, Vladimir Kushnarenko, Daniel Scharon, Pia Schmücker, Marcel Waldvogel und Stefan Wesner. „SARA-Service: Langzeitverfügbarkeit und Publikation von Softwareartefakten“ (Poster auf den E-Science-Tagen in Heidelberg März 2017). <https://doi.org/10.11588/heidok.00022887>.
- Arbeitskreis Open Science. „Helmholtz Open Science: Zugang zu und Nachnutzung von wissenschaftlicher Software.“ 2017. Zuletzt geprüft am 12.02.2018. <https://os.helmholtz.de/?id=2766>.
- „bwIDM | Förderiertes Identitätsmanagement der baden-württembergischen Hochschulen.“ Zuletzt geprüft am 13.02.2018. <https://www.bwidm.de/>.
- Clarivate Analytics. „Data Citation Index.“ Zuletzt geprüft am 15.02.2018. http://wokinfo.com/products_tools/multidisciplinary/dci/.
- Deutsche Forschungsgemeinschaft. *Sicherung guter wissenschaftlicher Praxis: Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft“*. Ergänzte Auflage. Weinheim: Wiley-VHC, 2013. <https://doi.org/10.1002/9783527679188.oth1>.
- Klaß, Christian. „Geocities geschlossen - und für die Nachwelt archiviert.“ *Golem.de*. Zuletzt geprüft am 12.02.2018. <https://www.golem.de/0910/70706.html>.

- Ministerium für Wissenschaft, Forschung und Kunst. „E-Science: Wissenschaft unter neuen Rahmenbedingungen.“ Zuletzt geprüft am 12.02.2018. <https://mwk.baden-wuerttemberg.de/de/forschung/forschungslandschaft/e-science/>.
- Smith, Arfon M., Daniel S. Katz, Kyle E. Niemeyer und FORCE11 Software Citation Working Group. „Software Citation Principles.“ *PeerJ Computer Science* 2 (2016): e86. <https://doi.org/10.7717/peerj-cs.86>.

Sozialwissenschaftliche Forschungsdaten langfristig sichern und zugänglich machen: Herausforderungen und Lösungsansätze

Anja Perry, *GESIS – Leibniz-Institut für Sozialwissenschaften, Köln*

Jonas Recker, *GESIS – Leibniz-Institut für Sozialwissenschaften, Köln*

Zusammenfassung:

Sozialwissenschaftlerinnen und -wissenschaftler sind gemäß der guten wissenschaftlichen Praxis dazu angehalten, den Forschungsprozess möglichst transparent zu halten, ihre Forschungsergebnisse reproduzierbar zu machen und im Sinne von „Open Science“ erhobene Daten für die Nachnutzung zur Verfügung zu stellen. Dies ist allerdings mit einigen Hürden verbunden. So müssen nachnutzbare Daten so dokumentiert werden, dass sie für Dritte auffindbar und verständlich sind. Gleichzeitig müssen personenbezogene Daten der Teilnehmenden ausreichend geschützt werden. Forschende sind hier auf Unterstützung angewiesen, um diesen Ansprüchen an die empirische Forschung gerecht zu werden. Dieser Artikel geht zunächst auf die Hürden bei der Bereitstellung von sozialwissenschaftlichen Forschungsdaten ein und stellt dann Dienstleistungen des GESIS Datenarchivs für Sozialwissenschaften vor, die Forschenden helfen, diese Hürden zu meistern.

Summary:

In accordance with the rules of good scientific practice, social scientists are encouraged to keep the research process as transparent as possible, to make their research results replicable, and – in the spirit of “Open Science” – to make collected data available for re-use. However, this is associated with a number of hurdles. For example, reusable data must be documented in such a way that it can be found and understood by third parties. At the same time, personal data of respondents must be adequately protected. Researchers need support to meet these demands on empirical research. The article first discusses the barriers to the provision of social science research data and afterwards presents special services of the GESIS Data Archive for Social Sciences, which help researchers to overcome these hurdles.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S106-122>

Autorenidentifikation: Anja Perry: ORCID <https://orcid.org/0000-0003-0574-9275>; Jonas

Recker: ORCID <https://orcid.org/0000-0001-9562-3339>

Schlagwörter: Sozialwissenschaften, Forschungsdaten, Datenmanagement, Archivierung

1. Einleitung

Die Anforderungen an Wissenschaftlerinnen und Wissenschaftler, möglichst umfassend Transparenz über den Forschungsprozess und die erzielten Ergebnisse herzustellen und damit Forschungsergebnisse reproduzierbar und nachvollziehbar zu machen, sind in den letzten Jahren stetig gestiegen. Enthielten schon die Regeln zur Sicherung der guten wissenschaftlichen Praxis der Deutschen

Forschungsgemeinschaft¹ Forderungen mit Blick auf eine Aufbewahrung von „Primärdaten“ inklusive einer Beschreibung der Erhebungsmethoden, so knüpfen Drittmittelgeber Förderzusagen heute zunehmend an das Vorliegen eines geeigneten Datenmanagementplans, welcher im Projektverlauf umgesetzt wird und idealerweise in einer Bereitstellung der Forschungsdaten für die Nachnutzung mündet. Das BMBF fordert in einzelnen Programmen beispielsweise, 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“.² Vergleichbare Forderungen oder mindestens Empfehlungen finden sich bei der DFG und der EU (Horizon 2020).³

Auch wissenschaftliche Zeitschriften beziehen die einer Publikation zugrundeliegenden Forschungsdaten immer häufiger in den Peer-Review-Prozess mit ein und empfehlen oder fordern die Veröffentlichung dieser Daten in geeigneten Systemen.⁴

Im Zuge dieser Entwicklung ist ein Bedarf an Forschungsdateninfrastrukturen entstanden, d.h. an Einrichtungen, die (fachspezifische) „Instrumente, Ressourcen und Dienstleistungen zur Verfügung [stellen]“.⁵ Dieser Bedarf wird mittlerweile auf nationaler und internationaler Ebene von einer Vielzahl an Initiativen und Projekten adressiert, in deren Rahmen disziplinspezifische und generische Angebote zum Management, der Publikation und der langfristigen Archivierung von Forschungsdaten entstehen. Hochschulen spielen hierbei neben disziplinspezifischen Infrastrukturanbietern eine wichtige Rolle: Sie vermitteln Forschenden Wissen und Kompetenzen für den Umgang mit Forschungsdaten; sie erarbeiten institutionelle Forschungsdatenpolicies und -leitlinien und setzen diese um; sie etablieren Beratungs- und Unterstützungsangebote für Forschende und stellen Systeme zur Veröffentlichung von Forschungsdaten bereit, häufig in enger Zusammenarbeit zwischen Bibliotheken, Rechenzentren und der Hochschulleitung. Beispiele hochschulischer Initiativen auf Landesebene sind die Projekte „bwFDM – Koordiniertes Forschungsdatenmanagement in Baden-Württemberg“ (abgeschlossen)

1 Vgl. Deutsche Forschungsgemeinschaft, *Sicherung guter wissenschaftlicher Praxis / Safeguarding Good Scientific Practice*. Denkschrift / Memorandum, 2013, Empfehlung 7, zuletzt geprüft am 09.03.2018, http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf.

2 Bundesministerium für Bildung und Forschung, „Bekanntmachung Richtlinie zur Förderung von Forschung zu ‚Digitalisierung im Bildungsbereich – Grundsatzfragen und Gelingensbedingungen‘“, Bundesanzeiger, Nr. AT 26.09.2017 B2 (2017), zuletzt geprüft am 14.02.2018, <https://www.bmbf.de/foerderungen/bekanntmachung-1420.html>.

3 Deutsche Forschungsgemeinschaft, *Umgang mit Forschungsdaten: DFG-Leitlinien zum Umgang mit Forschungsdaten*, 2015, zuletzt geprüft am 27.06.2018, http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/;

European Commission, *H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020*, 2016, zuletzt geprüft am 14.02.2018, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.

4 Beispiele sind PLOS (<http://journals.plos.org/plosone/s/data-availability>) und die Zeitschrift für Soziologie (<https://www.degruyter.com/view/j/zfsoz#callForPapersHeader>). Links zuletzt geprüft am 14.02.2018.

5 Wissenschaftsrat, *Bericht zur wissenschaftsgeleiteten Bewertung umfangreicher Forschungsinfrastruktur-vorhaben für die Nationale Roadmap*, 2017, 8, zuletzt geprüft am 14.02.2018, <https://www.wissenschaftsrat.de/download/archiv/6410-17.pdf>.

und „Hessische Forschungsdateninfrastrukturen – HeFDI“; auch in Nordrhein-Westfalen besteht seit 2016 eine einrichtungsübergreifende Kooperation.⁶

Eine Herausforderung, vor der solche hochschulischen Angebote stehen, ist, dass sie eine große Bandbreite an wissenschaftlichen Fachrichtungen mit ganz unterschiedlichen Voraussetzungen und Anforderungen bezüglich Beratung oder technischer Infrastruktur bedienen müssen. In der empirischen Sozialforschung spielen hier insbesondere datenschutzrechtliche Fragen eine Rolle, da häufig so genannte personenbezogene Daten erhoben und verarbeitet werden. Auch der Erhalt der Interpretierbarkeit der Daten, der maßgeblich von der Verfügbarkeit ausreichender Kontextinformationen zur Datenerhebung und -verarbeitung abhängt, folgt disziplinspezifischen Praktiken und Standards.

Oft ist es nicht möglich, im Rahmen einer generischen Infrastruktur allen fachspezifischen Anforderungen umfassend und in ausreichender Tiefe gerecht zu werden. Hierbei können disziplinspezifische Infrastrukturen wie das GESIS – Leibniz-Institut für Sozialwissenschaften wichtige Aufgaben übernehmen, indem sie fachbezogene Unterstützungs- und Lösungsangebote zur langfristigen Sicherung und Bereitstellung von Forschungsdaten bieten und damit generische Angebote ergänzen. Nach einer kurzen Beschreibung, was unter sozialwissenschaftlichen Forschungsdaten verstanden wird, soll im folgenden Beitrag vor diesem Hintergrund ein Bewusstsein für Herausforderungen geschaffen werden, die bei der langfristigen Sicherung und Bereitstellung sozialwissenschaftlicher Forschungsdaten zu beachten sind. Anschließend werden beispielhaft Unterstützungs- und Lösungsangebote von GESIS – Leibniz-Institut für Sozialwissenschaften skizziert.

2. Was sind sozialwissenschaftliche Daten?

Forschende in den Sozialwissenschaften arbeiten mit einer großen Bandbreite unterschiedlicher Arten von Daten. Eine typische Unterscheidung wird dabei zwischen quantitativen und qualitativen Daten getroffen. Quantitative sozialwissenschaftliche Daten sind häufig Umfragedaten. Das können Meinungsumfragen sein, wie das Politbarometer,⁷ Vergleichsstudien, wie das Programme for the International Comparison of Adult Competencies (PIAAC) zum internationalen Vergleich der Kompetenzen Erwachsener,⁸ oder auch Langzeitbefragungen, wie das Beziehungs- und Familienpanel pairfam.⁹ Auch Prozess- oder Registerdaten, wie z.B. die Verwaltungsdaten der sozialversicherungspflichtig Beschäftigten bei der Bundesagentur für Arbeit, werden für Auswertungen verwendet. Aus dieser Bandbreite an Daten ziehen Sozialwissenschaftlerinnen und Sozialwissenschaftler Rückschlüsse über die Gesellschaft und geben Handlungsempfehlungen an die Politik.

6 „bwFDM - Koordiniertes Forschungsdatenmanagement in Baden-Württemberg,“ zuletzt geprüft am 14.02.2018, <https://bwfdm.scc.kit.edu/>; „Hessische Forschungsdateninfrastrukturen - HeFDI,“ zuletzt geprüft am 14.02.2018, <https://www.uni-marburg.de/projekte/forschungsdaten/projekt>; Constanze Curdt et al., „Herausforderung Forschungsdatenmanagement – Unterstützung der Hochschulen durch eine einrichtungsübergreifende Kooperation in NRW,“ in *E-Science-Tage 2017: Forschungsdaten managen*, hrsg. Jonas Kratzke und Vincent Heuveline (Heidelberg: heiBOOKS, 2017), <https://doi.org/10.11588/heibooks.285.377>.

7 <https://doi.org/10.4232/1.12824>.

8 <http://www.oecd.org/skills/piaac/>.

9 <https://doi.org/10.4232/pairfam.5678.8.0.0>.

Die Daten sind in der Regel in einer Datenmatrix organisiert und stark strukturiert. In dieser Datenmatrix sind die befragten Personen in Zeilen angeordnet (z.B. eine Zeile pro befragter Person) und die einzelnen Variablen spaltenweise angeordnet. Oftmals liegen diese Daten in proprietären Formaten, wie den gängigen Statistikformaten SPSS und Stata, oder aber auch als Tabelle (MS Excel) oder im Textformat (csv) vor (s. Abbildung 1).

	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
1	20001 1.0.0 (2017-07-01)		1	2	5	2	1	2	99	1	1
2	20002 1.0.0 (2017-07-01)		1	2	5	2	5	2	2	1	1
3	20003 1.0.0 (2017-07-01)		1	2	17	1	8	1	11	1	1
4	20004 1.0.0 (2017-07-01)		1	2	5	2	4	2	29	1	2
5	20005 1.0.0 (2017-07-01)		1	2	15	2	5	3	35	1	1
6	20006 1.0.0 (2017-07-01)		1	2	8	2	4	46	99	1	5
7	20007 1.0.0 (2017-07-01)		1	2	16	1	1	2	7	1	2
8	20008 1.0.0 (2017-07-01)		1	2	15	1	5	2	24	1	6
9	20009 1.0.0 (2017-07-01)		1	2	17	1	3	2	99	2	0
10	20010 1.0.0 (2017-07-01)		1	2	14	1	4	2	29	1	2
11	20011 1.0.0 (2017-07-01)		1	2	5	2	8	2	8	1	99
12	20012 1.0.0 (2017-07-01)		1	2	17	1	1	2	99	1	3
13	20013 1.0.0 (2017-07-01)		1	2	6	2	2	38	2	1	99
14	20014 1.0.0 (2017-07-01)		1	2	15	1	6	2	35	2	99
15	20015 1.0.0 (2017-07-01)		1	2	2	2	8	2	14	1	2
16	20016 1.0.0 (2017-07-01)		1	2	12	1	8	2	36	1	3
17	20017 1.0.0 (2017-07-01)		1	2	5	2	7	2	32	1	2
18	20018 1.0.0 (2017-07-01)		1	2	15	1	2	2	3	1	99
19	20019 1.0.0 (2017-07-01)		1	2	16	1	7	2	40	2	0
20	20020 1.0.0 (2017-07-01)		1	2	1	2	1	2	99	1	1

Abbildung 1: Beispiel einer Datenmatrix in SPSS

Qualitative Daten hingegen sind nicht-numerische Daten, die häufig schriftlich (z.B. Interviewtranskripte, Beobachtungsprotokolle, Feldtagebücher oder Texte in Tageszeitungen) oder als Audio- oder Videoaufzeichnungen (z.B. von Gruppendiskussionen oder Beobachtungsstudien) vorliegen. Diese sind, im Gegensatz zu quantitativen Daten, daher oft nicht stark strukturiert.

3. Der Übergang sozialwissenschaftlicher Daten in die Langzeitsicherung: Herausforderungen und Lösungen

3.1. Die Interpretierbarkeit sicherstellen

Daten die anderen Forschenden zur Nutzung bereitgestellt werden, müssen für diese langfristig interpretierbar sein. Dies erfordert eine gründliche Aufbereitung und Dokumentation der Daten, da diese nicht für sich selbst sprechen. Es sind also zusätzliche Informationen nötig, die die Daten beschreiben und damit interpretierbar machen. Diese werden den Daten in Form von strukturierten oder unstrukturierten Metadaten („Dokumentation“) zur Seite gestellt und dokumentieren beispielsweise, welche Personen unter welchen Bedingungen befragt wurden.

Im Fall von quantitativen Daten unterscheidet man in Dokumentation auf Studienebene und auf Variablenebene. Die Dokumentation auf Studienebene umfasst, wo zutreffend, den Fragebogen, einen Methodenbericht sowie weitere Dokumente, die zum Verstehen der Studie als Ganzes wichtig sein könnten. Hier können Besonderheiten bei der Erhebung eine Rolle spielen. Die Einstellungen zur Atomkraft können sich beispielsweise kurz nach dem Reaktorunfall von Fukushima stark verändert haben und in Befragungen zu diesem Thema zu starken Trendänderungen führen. Der Fragebogen

erlaubt, die Originalfrage zu betrachten und Rückschlüsse auf das Antwortverhalten zu ziehen. Der Methodenbericht gibt einen Überblick über Erhebungsdesign, Erhebungsprozess und Datenverarbeitung.¹⁰ Zum Beispiel ist die Antwortrate ein entscheidendes Merkmal, um einschätzen zu können, wie repräsentativ die Erhebung ist. Beide Dokumente erlauben somit Dritten, die Qualität der Erhebung und ggf. die Nutzbarkeit der Daten für ihre eigenen Forschungsinteressen zu bewerten.

Auf Variablenebene fokussiert die Dokumentation auf die Beschreibung, die Struktur und den Inhalt der Daten. Ziel ist es, den Nutzenden einen schnellen Überblick über die Daten zu geben. Zentral sind hierbei die Dokumentation der Variablen im Datensatz selbst (durch Variablenname und -label) sowie das Codebook. Darin sind alle Variablen mit Namen, Label, Format, Ausprägungen (inkl. der Kennzeichnung fehlender Werte) und zugrundeliegenden Konzepten aufgelistet. Es stellt damit die Verbindung zwischen Variablen im Datensatz und Originalvariable im Fragebogen her.¹¹

Im Fall von qualitativen Daten müssen ebenfalls auf Studienebene Kontextinformationen, wie Ort der Erhebung und Zeitgeschehnisse, dokumentiert werden. Außerdem sollten beispielsweise im Fall von Interviews der Interviewleitfaden, ursprüngliche Fragenformulierungen, Anweisungen an die interviewenden Personen sowie zusätzliche Erklärungen nachgehalten werden. Im Fall von Ethnographie und (teilnehmenden) Beobachtungsstudien sind dies z.B. Feldtagebücher oder Auszüge daraus. Auf Fallebene muss erkennbar sein, wie verschiedene Dateien (z.B. Audiomitschnitt und Transkription) zueinander stehen, wer die Interviews oder Videoaufnahmen transkribiert hat und welche Situationsmerkmale in den jeweiligen Interviews eine Rolle gespielt haben.¹²

Um eine gute Aufbereitung und Dokumentation sicherzustellen, wird empfohlen, von Beginn des Forschungsprojekts an einen Datenmanagementplan zu verfassen.¹³ Dieser legt Verantwortliche für bestimmte Aufgaben im Projekt fest und definiert Abläufe für diese Aufgaben. Er wird im Laufe des Projekts immer wieder angepasst, um neu entstehenden Problemen und Aufgaben im Projekt gerecht zu werden.

10 Vgl. Oliver Watteler, *Erstellung von Methodenberichten für die Archivierung von Forschungsdaten*, 2010, zuletzt geprüft am 14.02.2018, https://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/Aufbau_Methodenbericht_v1_2010-07.pdf; Verbund Forschungsdaten Bildung, *Hinweise zur Erstellung eines Methoden- / Feldberichts*, fdbinfo Nr. 3, 2015, zuletzt geprüft am 14.02.2018, https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdbinfo_3.pdf.

11 Vgl. z.B. Uwe Jensen, „Leitlinien zum Management von Forschungsdaten. Sozialwissenschaftliche Umfragedaten,“ *GESIS Technical Reports*, Nr. 7 (2012), <http://nbn-resolving.de/urn:nbn:de:O168-ssoar-320650>; Karoline Harzenetter, „Metadata documentation at the variable level,“ in *Data Processing and Documentation: Generating High Quality Research Data in Quantitative Social Science Research* (GESIS Papers), hrsg. Sebastian Netscher und Christina Eder (Köln: GESIS - Leibniz-Institut für Sozialwissenschaften, in Vorbereitung).

12 Vgl. Cabrera und Haake, „Das Qualiservice Metadatenschema. Version 1.0,“ *Qualiservice Technical Reports*, Nr. 13 (2013), <http://nbn-resolving.de/urn:nbn:de:gbv:46-00103310-19>.

13 Vgl. z.B. Ulf Jakobsson, Ricarda Braukmann und Malin Lundgren, „Expert tour guide on Data Management. Chapter 1: Plan,“ 2017, zuletzt geprüft am 14.02.2018, <http://www.cessda.eu/DMGuide>; Rat für Sozial- und Wirtschaftsdaten, *Forschungsdatenmanagement in den Sozial-, Verhaltens- und Wirtschaftswissenschaften. Orientierungshilfen für die Beantragung und Begutachtung datengenerierender und datennutzender Forschungsprojekte*, RatSWD Output Series 3 (Berlin, 2016), <https://doi.org/10.17620/02671.7>.

3.2. Ethisch-rechtliche Anforderungen

Die Erhebung, Aufbereitung und Archivierung von Umfrage- und Registerdaten sowie qualitativen Daten betrifft immer auch personenbezogene Daten. Personenbezogene Daten sind „alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person [...] beziehen“.¹⁴ Diese zu schützen, um Teilnehmende vor negativen Folgen der Forschung sowie einer kommerziellen Nutzung der Daten zu bewahren, ist ein Grundrecht der Teilnehmenden nach der Europäischen Menschenrechtskonvention.¹⁵ Aus diesem Grund müssen sich Sozialwissenschaftlerinnen und Sozialwissenschaftler möglichst frühzeitig auch mit dem Datenschutz auseinandersetzen. Dies sollte schon bei der Fragebogenerstellung geschehen, indem im Sinne der Datensparsamkeit nur Fragen gestellt werden, die tatsächlich für die Forschungsfrage relevant sind (Datenminimierung).¹⁶

Zudem muss für die Erhebung personenbezogener Daten eine Einverständniserklärung eingeholt werden.¹⁷ In Tabelle 1 werden die wichtigsten Bestandteile der informierten Einverständniserklärung zusammengefasst. In aller Regel wird das informierte Einverständnis als freiwillige aktive Zustimmung zur Teilnahme vor Beginn der Datenerhebung eingeholt. In Ausnahmefällen, z.B. wenn der Zweck der Forschung den Teilnehmenden nicht vorab bekannt sein darf, kann diese auch im Nachhinein eingeholt werden.¹⁸ Der Informationsteil mit den in Tabelle 1 genannten Bestandteilen sollte bei den Teilnehmenden verbleiben. Das deutsche Recht verlangte bisher eine schriftliche Einverständniserklärung. Laut der neuen europäischen Datenschutzgrundverordnung (DSGVO) ist eine schriftliche Einverständniserklärung seit dem 25.05.2018 nicht mehr zwingend nötig. Jedoch müssen Forschende nachweisen können, dass sie die informierte Einwilligung eingeholt haben.¹⁹ Dafür ist weiterhin eine schriftliche Einwilligung am besten geeignet. Bei bestimmten Gruppen von Teilnehmenden kann und muss man allerdings davon abweichen. Zum Beispiel können Personen ohne ausreichende Lese-/ Schreibkompetenz die informierte Einwilligung nur mündlich geben.

14 Art. 4 „Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (DSGV)“ (2016), zuletzt geprüft am 14.02.2018, <http://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32016R0679&from=DE>.

15 Vgl. Art. 8 Abs. 1 „Europäische Menschenrechtskonvention (Konvention zum Schutze der Menschenrechte und Grundfreiheiten) vom 04.11.1950,“ zuletzt geprüft am 14.02.2018, <https://dejure.org/gesetze/MRK>.

16 Vgl. Art. 5 Abs. 1 DSGVO.

17 Vgl. Art. 13 Abs. 1 ebd.

18 Vgl. Deutsche Gesellschaft für Soziologie und Berufsverband Deutscher Soziologinnen und Soziologen, „Ethik-Kodex der Deutschen Gesellschaft für Soziologie (DGS) und des Berufsverbandes Deutscher Soziologinnen und Soziologen (BDS),“ 2017, §2 Abs. 3, zuletzt geprüft am 10.03.2018, <http://www.soziologie.de/de/die-dgs/ethik/ethik-kodex.html>.

19 Vgl. Art. 7 Abs. 1 ebd.; für eine Zusammenfassung der für die Forschung relevante Neuerungen durch die DSGVO vgl. Katrin Schaar, *Was hat die Wissenschaft beim Datenschutz künftig zu beachten? Allgemeine und spezifische Änderungen beim Datenschutz im Wissenschaftsbereich durch die neue Europäische Datenschutzgrundverordnung*, Working Paper Series des Rates für Sozial- und Wirtschaftsdaten, Nr. 257 (2016), <https://doi.org/10.17620/02671.19>.

Tabelle 1: Bestandteile einer informierten Einverständniserklärung

Inhaltlicher Informationsteil	<ul style="list-style-type: none">• Projekttitel• Zweck der Studie• Namen der beteiligten Forschenden• Erhebung personenbezogener Daten• Datennutzung im Projekt und über das Projekt hinaus• Aufbewahrungsfristen für die Daten• Einsichtsrechte und Weitergabe an Dritte (z.B. Archivierung und Nachnutzung)• Schutz persönlicher Angaben (z.B. durch Pseudonymisierung)• „in präziser, transparenter, verständlicher und leicht zugänglicher Form in einer klaren und einfachen Sprache“²⁰
Rechtlicher Informationsteil	<ul style="list-style-type: none">• Freiwilligkeit der Teilnahme• Informationsrecht• Recht auf Rücknahme der Einwilligung zu jeder Zeit• Recht auf Einsichtnahme in die persönlichen Daten• Recht auf Verbesserung der persönlichen Daten• Recht auf Löschung der persönlichen Daten
Einwilligungs- erklärung	<ul style="list-style-type: none">• Besonderheiten der Befragung (z.B. Videoaufzeichnung) müssen abgedeckt sein• Stabile Kontaktinformation für Rückfragen und Wahrnehmung der zugesicherten Rechte• Unterschrift

Auch für die Archivierung und Nachnutzung der erhobenen Daten ist die informierte Einwilligung wichtig. Fehlt diese, fehlt die rechtliche Grundlage für die Datenerhebung und damit auch für die Aufbewahrung und Nachnutzung der Daten. Aber auch eine unglückliche Formulierung, wie z.B. die Zusicherung, dass die Daten nur im Forschungsprojekt verwendet werden, schließen eine Archivierung und eine Nutzung über den Zweck des Forschungsprojektes hinaus aus. Die neue DSGVO verlangt präzise und transparente Informationen über die Erhebung, Weiterverarbeitung und Aufbewahrung der Daten.²¹

Forschungsdaten müssen für die Weiterverarbeitung anonymisiert werden. Hier galt bislang in Deutschland der Grundsatz der faktischen Anonymisierung.²² Dabei dürfen die Einzelangaben „nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer

²⁰ Art. 12 Abs. 1 Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (DSGV).

²¹ Vgl. Schaar, *Was hat die Wissenschaft beim Datenschutz künftig zu beachten?*

²² Vgl. Heike Wirth, „Die faktische Anonymität von Mikrodaten: Ergebnisse und Konsequenzen eines Forschungsprojektes,“ ZUMA Nachrichten 16, Nr. 30 (1992): 7-65, <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-209679>.

bestimmten oder bestimmbar natürlichen Person zugeordnet werden können“.²³ Nach einer Stellungnahme der Artikel-29-Datenschutzgruppe greift die neue DSGVO die faktische Anonymisierung auf und verlangt bei der Anonymisierung oder Pseudonymisierung eine Abwägung, ob eine Person bestimmbar ist. Bei dieser Abwägung müssen Kosten der Identifizierung, erforderlicher Zeitaufwand sowie verfügbare Technologien und technologische Entwicklungen berücksichtigt werden.²⁴

Bei der Pseudonymisierung wird in direkte und indirekte Identifikatoren unterschieden. Wie diese definiert sind und wie mit ihnen umgegangen wird, um die erhobenen Daten zu anonymisieren, ist in Tabelle 2 dargestellt. Alle Maßnahmen zur Pseudonymisierung müssen sorgfältig für die Datennutzer/innen dokumentiert werden, um falsche Rückschlüsse aus den Daten zu vermeiden.

Tabelle 2: Direkte und indirekte Identifikatoren und Pseudonymisierungsstrategien²⁵

Art des Identifikators	Definition	Pseudonymisierungsstrategie
Direkt	Angaben, die Teilnehmende direkt identifizieren, wie z.B. Name und Kontaktdaten, Bild oder Stimme bei audiovisuellen Medien	<ul style="list-style-type: none"> • Schnellstmögliche, dauerhafte Löschung dieser Angaben nach der Befragung • Verpixelung und/oder Verzerrung von Bild-/Tondokumenten • Verwendung von Substituten, wie z.B. „Schüler 1“ • Erfolgt eine Wiederbefragung, müssen sie getrennt von den erhobenen Daten sicher und unzugänglich für Dritte aufbewahrt werden

23 § 3 Abs. 6 „Bundesdatenschutzgesetz (BDSG)“, zuletzt geprüft am 27.06.2018, https://dejure.org/gesetze/BDSG_a.F./3.html.

24 Vgl. Erwägungsgrund Nr. 26 DSGVO; Article 29 Data Protection Working Party, „Opinion 05/2014 on Anonymisation Techniques“, 2014, zuletzt geprüft am 16.02.2018, http://collections.internetmemory.org/haeu/20140424090503/http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

25 Vgl. Heike Wirth, *Anonymisierung des Mikrozensuspanels im Kontext der Bereitstellung als Scientific-Use-File*, 2006, zuletzt geprüft am 14.02.2018, <https://www.destatis.de/DE/Methoden/Methodenpapiere/Mikrozensus/Arbeitspapiere/Arbeitspapier11.html>; Oliver Watteler und Katharina E. Kinder-Kurlanda, „Anonymisierung und sicherer Umgang mit Forschungsdaten in der empirischen Sozialforschung,“ *Datenschutz und Datensicherheit : DuD ; Recht und Sicherheit in Informationsverarbeitung und Kommunikation* 39, Nr. 8 (2015): 515–19, <https://doi.org/10.1007/s11623-015-0462-0>; Alexia Meyermann und Maïke Porzelt, *Hinweise zur Anonymisierung von qualitativen Daten*, forschungsdaten bildung informiert, Nr. 1, 2014, zuletzt geprüft am 14.02.2018, https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert-nr-1.pdf; Marcus Eisentraut, „Data anonymization,“ in *Data Processing and Documentation: Generating High Quality Research Data in Quantitative Social Science Research* (GESIS Papers), hrsg. von Sebastian Netscher und Christina Eder (Köln: GESIS - Leibniz-Institut für Sozialwissenschaften, in Vorbereitung).

Indirekt	Informationen, durch die in Kombination miteinander Personen identifiziert werden können, z.B. weil sie sehr kleinteilig sind (wie der Wohnort) oder Merkmale, die nur sehr selten in der Grundgesamtheit auftreten (wie ein bestimmtes Herkunftsland)	<ul style="list-style-type: none">• Vergrößerung regionaler Informationen• Zusammenfassen von Merkmalen in Kategorien• Aggregieren von Informationen
-----------------	--	--

Nicht zuletzt müssen personenbezogene Daten auch vor unautorisiertem Zugriff durch Dritte geschützt werden.²⁶ Dies kann durch Zugangskontrolle über ein sicheres Passwort oder durch Verschlüsselung der Daten erfolgen. Daten die nicht mehr verwendet werden, sollten schnellstmöglich dauerhaft gelöscht werden.

4. Sicherung und Bereitstellung von Umfragedaten am GESIS Datenarchiv für Sozialwissenschaften

Gegründet im Jahr 1960 als Zentralarchiv für empirische Sozialforschung ist das Datenarchiv für Sozialwissenschaften heute eine Abteilung von GESIS – Leibniz-Institut für Sozialwissenschaften, der größten europäischen Infrastruktureinrichtung für die Sozialwissenschaften (www.gesis.org). GESIS unterstützt Wissenschaftlerinnen und Wissenschaftler, die mit Methoden der quantitativen Sozialforschung arbeiten, in allen Phasen des Forschungsprozesses. Zu den forschungsbasierten Angeboten und Dienstleistungen gehört neben der methodischen Beratung zur Studienplanung und -implementierung, der Durchführung kognitiver Pretests oder der Ziehung von Telefonstichproben zum Beispiel auch die Unterstützung von Forschenden und Forschungseinrichtungen beim Management und der langfristigen Sicherung sowie der Bereitstellung quantitativer sozialwissenschaftlicher Forschungsdaten.²⁷ So stehen im GESIS Datenarchiv über 5500 Studien für die Nachnutzung zur Verfügung und können in der Regel nach einer Registrierung im Datenbestandskatalog DBK direkt für die eigene Forschung und Lehre heruntergeladen werden.²⁸ Forschungsdatenzentren wie z.B. das FDZ ALLBUS oder das FDZ Internationale Studien sorgen für eine added-value Aufbereitung und Bereitstellung von Daten aus großen Studienprogrammen.²⁹

4.1. Übernahme von quantitativen Forschungsdaten in das GESIS Datenarchiv für Sozialwissenschaften

Um den oben skizzierten Herausforderungen bei der langfristigen Sicherung und Bereitstellung von sozialwissenschaftlichen Forschungsdaten für die Nutzung zu begegnen, führt das GESIS Datenarchiv für Sozialwissenschaften bei der Übernahme solcher Daten eine Reihe von Überprüfungen durch. Diese sind geleitet von folgenden zentralen Fragestellungen:

²⁶ Vgl. Art. 5 Abs. 1 DSGVO.

²⁷ Weitere Informationen zum Angebot von GESIS sind unter <https://www.gesis.org/angebot/> abrufbar.

²⁸ Datenbestandskatalog, <https://dbk.gesis.org/dbksearch?db=d>, zuletzt geprüft am 14.02.2018.

²⁹ GESIS Forschungsdatenzentren, <https://www.gesis.org/institut/forschungsdatenzentren/>, zuletzt geprüft am 14.02.2018.

- Liegen ausreichend Informationen vor, sodass Dritte die eingereichten Forschungsdaten verstehen und nutzen können?
- Sind die Forschungsdaten von guter Qualität (d.h. z.B. fehlerfrei aufbereitet)?
- Gibt es technische Probleme bei der Nutzung der eingereichten Forschungsdaten und Dokumentation?
- Bestehen rechtliche Probleme im Hinblick auf den Datenschutz oder das Urheberrecht?

Um diese Fragen zu beantworten, führen erfahrene Kuratorinnen und Kuratoren umfassende Überprüfungen durch (siehe Tabelle 3).

Tabelle 3: Überprüfungen bei Dateneingang im GESIS Datenarchiv (Auswahl)

technisch	<ul style="list-style-type: none">• Lesbarkeit und Virenfreiheit aller Dateien
studienbezogen	<ul style="list-style-type: none">• Vollständigkeit der eingereichten Dokumente (maschinenlesbarer Datensatz, Codebuch, Erhebungsinstrument, Methodenbericht);• die Einhaltung des Datenschutzes (sind Daten ausreichend anonymisiert?) sowie des Urheberrechts (wurden z.B. urheberrechtlich geschützte Instrumente verwendet?).
datensatzbezogen	<ul style="list-style-type: none">• Vollständigkeit der Fälle• Vollständigkeit und Korrektheit der Variablen, Variablennamen und Variablenlabel (auch: passen die Variablen zum Fragebogen?)• Gewichte• Plausibilität (u.a. Verteilung nach Geschlecht und Alter)• undokumentierte Werte (so genannte „wild codes“), fehlende Werte

Sofern hier Probleme deutlich werden, wird in Absprache mit den datengebenden Forscherinnen und Forschern versucht, diese auszuräumen – z.B. indem die Pseudonymisierung angepasst wird. Sollten Inkonsistenzen auffallen, die in einer fehlerhaften Verarbeitung der Daten begründet sind, erfolgt in Rücksprache mit den datengebenden Personen eine Korrektur. Inkonsistenzen die sich nicht korrigieren lassen – weil sie z.B. in einem verzerrten Antwortverhalten begründet sind – werden dokumentiert, damit Forschende, die die Daten nachnutzen, darüber informiert sind.³⁰

30 Vgl. Hannah Schwarz, „Data consistency,“ in *Data Processing and Documentation: Generating High Quality Research Data in Quantitative Social Science Research* (GESIS Papers), hrsg. von Sebastian Netscher und Christina Eder (Köln: GESIS - Leibniz-Institut für Sozialwissenschaften, in Vorbereitung).

4.2. Disziplinspezifische strukturierte Metadaten für Auffindbarkeit und Zitierbarkeit

Wie oben geschildert, wird den Forschungsdaten mit der sogenannten „Dokumentation“ relevante Kontextinformation in unstrukturierter Form zur Seite gestellt. Diese gewährleistet die Interpretierbarkeit der Daten, ist aber aufgrund der mangelnden Maschinenlesbarkeit nur bedingt geeignet, die Auffindbarkeit der Daten einerseits und die schnelle Beurteilung der Relevanz der Daten für die eigene Forschung andererseits zu ermöglichen. Diese Funktion erfüllen strukturierte Metadaten, die unter Verwendung disziplinspezifischer Standards erstellt werden – von den datengebenden Wissenschaftlerinnen und Wissenschaftlern selbst bei Einreichung und/oder von den Kuratorinnen und Kuratoren, die die Daten dann ins Archiv übernehmen. Das GESIS Datenarchiv für Sozialwissenschaften verwendet dafür ein eigenes Metadatenschema.³¹ Es bestehen Mappings zu disziplinspezifischen Standards, insbesondere dem *dalra* Metadatenschema für die Registrierung von DOIs in den Sozial- und Wirtschaftswissenschaften sowie den XML-basierten Schemata DDI-Codebook und DDI-Lifecycle der Data Documentation Initiative (<http://www.ddialliance.org>).³² Diese ermöglichen eine differenzierte Beschreibung der Forschungsdaten, einschließlich der Erhebung und Aufbereitung, und machen es damit einfacher, Forschungsdaten für die Nutzung in der eigenen Forschung zu finden und deren Relevanz zu beurteilen. Die vom GESIS Datenarchiv für Sozialwissenschaften vorgenommene Registrierung der Forschungsdaten bei *dalra*, der DOI-Registrierungsagentur für Sozial- und Wirtschaftsdaten, sichert darüber hinaus auch die eindeutige Identifizierbarkeit und Zitierbarkeit der Daten.

4.3. Ausgewählte Unterstützungsangebote

GESIS unterstützt Forschende, die mit sozialwissenschaftlichen Methoden arbeiten, beim Management und der langfristigen Sicherung und Bereitstellung ihrer Forschungsdaten. Dies geschieht über ein Schulungs- und Beratungsangebot sowie über die Bereitstellung einer Infrastruktur für die Veröffentlichung und Archivierung von Daten. So führt GESIS einführende Workshops in Themen des Forschungsdatenmanagements für Sozialwissenschaftlerinnen und wissenschaftler durch und unterstützt bei Fragen zur informierten Einwilligung oder der Pseudonymisierung von quantitativen Forschungsdaten.³³

Im Bereich der unterstützenden Infrastruktur bietet das GESIS Datenarchiv für Sozialwissenschaften mit dem Repository *datorium* eine Lösung an, die es Forschenden ermöglicht, ihre Forschungsdaten und zugehörige Dokumentation eigenständig mit Metadaten zu beschreiben und anderen unter standardisierten Lizenzbedingungen zur Nutzung zur Verfügung zu stellen. Zur Unterstützung von Institutionen, die selbst kein eigenes (disziplinspezifisches) Repository aufsetzen wollen, bietet das GESIS Datenarchiv für Sozialwissenschaften *SowiDataNet* als kostenpflichtige Business-to-business-Lösung

31 Wolfgang Zenk-Möltgen und Norma Habel, *Der GESIS Datenbestandskatalog und sein Metadatenschema. Version 1.8*, GESIS Technical Reports, Nr. 1 (2012), <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-292372>.

32 Ute Koch et al., *dalra Metadata Schema Documentation for the Publication and Citation of Social and Economic Data Version 4.0*, GESIS Papers, Nr. 25 (2017), <https://doi.org/10.4232/10.mdsdoc.4.0>; Esra Akdeniz und Wolfgang Zenk-Möltgen, *DDI-Lifecycle im Datenarchiv: das Metadatenschema für die Dokumentation in verschiedenen Softwaresystemen*, GESIS Papers, Nr. 2 (2017), <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-50354-9>.

33 <https://www.gesis.org/angebot/>.

an. Bei beiden Lösungen steht das GESIS Datenarchiv den Forschenden und kuratierenden Institutsmitgliedern als Ansprechpartner bei inhaltlichen und technischen Fragen zur Verfügung (siehe Tabelle 4). Weitere Angebote des GESIS Datenarchivs für Sozialwissenschaften umfassen z.B. die Langzeitarchivierung (u.a. mit Migration der Daten in andere Dateiformate) und die added-value Aufbereitung von Forschungsdaten (z.B. Erschließung auf Variableneben).

Tabelle 4: Überblick datorium und SowiDataNet

	datorium	SowiDataNet
Zielgruppe	Kleine Projekte, Einzelforscher/innen	Forschungseinrichtungen ohne eigenes Repository
Workflow	<ul style="list-style-type: none"> • Beschreibung mit Metadaten und Upload der Dateien durch die Forschenden; • Überprüfung von Daten und Dokumentation durch GESIS Datenkurator/in; • DOI-Registrierung und Veröffentlichung. 	<ul style="list-style-type: none"> • Beschreibung mit Metadaten und Upload der Dateien durch Institutsmitglieder (Forschende oder Kurator/Kuratorin); • Zweistufige Überprüfung der Daten durch Institution und GESIS Datenkurator/in; • DOI-Registrierung und Veröffentlichung.
Archivierung	Bitstream Preservation, d.h. Dateien werden im ursprünglichen Format langfristig erhalten	
URL	https://datorium.gesis.org/xmlui/	http://www.sowidatanet.de/xmlui

5. Sicherung und Bereitstellung qualitativer Forschungsdaten

Das GESIS Datenarchiv für Sozialwissenschaften bietet einen umfassenden Service für die Archivierung und Bereitstellung quantitativer Forschungsdaten. Für die weniger standardisierten qualitativen Forschungsdaten existiert bisher ein solch umfassendes Angebot noch nicht, obwohl auch hier die Nachnutzung der Daten eine wichtige Rolle spielt.³⁴ Daher wird in jüngster Zeit der Ruf nach einer Forschungsinfrastruktur auch für qualitative Daten laut.³⁵ Zwei themenspezifische Angebote

34 Vgl. Denis Huschka und Claudia Oellers, „Einführung: Warum qualitative Daten und ihre Sekundäranalyse wichtig sind,“ in *Forschungsinfrastrukturen für die qualitative Sozialforschung*, hrsg. Denis Huschka et al. (Berlin: Scivero, 2013), 9-18, zuletzt geprüft am 14.02.2018, https://www.germandataforum.org/dl/downloads/forschungsinfrastrukturen_qualitative_sozialforschung.pdf.

35 Vgl. Hubert Knoblauch, „Einigen Anforderungen an forschungsinfrastrukturen aus der Sicht der qualitativen Forschung,“ in *Forschungsinfrastrukturen für die qualitative Sozialforschung*, hrsg. Denis Huschka et al. (Berlin: Scivero, 2013), 27-34, zuletzt geprüft am 14.02.2018, https://www.germandataforum.org/dl/downloads/forschungsinfrastrukturen_qualitative_sozialforschung.pdf.

existieren bereits, die hier erwähnt werden sollen. Zum einen unterstützt der Qualiservice der Universität Bremen,³⁶ ausgehend vom Archiv für Lebenslaufforschung, die Archivierung von qualitativen Interviewstudien. Im Bereich der Bildungsforschung archiviert das FDZ Bildung am Deutschen Institut für Internationale Pädagogische Forschung³⁷ qualitative Beobachtungs- und Interviewdaten und stellt diese in anonymisierter Form für die wissenschaftliche Nachnutzung zur Verfügung.

6. Schlussfolgerungen

Zur Sicherung der guten wissenschaftlichen Praxis sollen Wissenschaftlerinnen und Wissenschaftler den Forschungsprozess möglichst transparent gestalten und die erhobenen Daten und erzielten Ergebnisse reproduzierbar und nachvollziehbar machen. Im Sinne von „Open Science“ sollen sie zudem eine möglichst freie und breite Nachnutzung ihrer Daten und Ergebnisse ermöglichen. Dies stellt Forschende aber auch vor große Hürden, da erhobene Daten möglichst für eine Nachnutzung dokumentiert und bereitgestellt werden sollen und gleichzeitig die Persönlichkeitsrechte der Teilnehmenden geschützt werden müssen.

Um Forschende hierbei zu unterstützen, bietet das GESIS-Datenarchiv für Sozialwissenschaften umfassende Informations- und Serviceangebote rund um die Archivierung quantitativer Forschungsdaten. Im Bereich der qualitativen Sozialforschung wurden mit dem Angebot von Qualiservice und dem FDZ Bildung bereits erste Grundsteine für eine unterstützende Infrastruktur gelegt. Jedoch muss sichergestellt werden, dass diese Angebote die Forschenden auch erreichen. Vor diesem Hintergrund erscheint es sinnvoll, die bislang noch nicht umfassend ausgeprägte Verzahnung generischer, insbesondere hochschulischer Angebote mit disziplinspezifischen Diensten zu stärken. Im großen Maßstab spielen diese Überlegungen nicht zuletzt in den aktuellen Planungen zur Nationalen Forschungsdateninfrastruktur (NFDI) eine Rolle, die sich wesentlich dadurch auszeichnen soll, dass generische und disziplinspezifische Angebote sich ergänzen, um unnötige Redundanzen zu vermeiden.³⁸ Im „Kleinen“ kann dies bedeuten, die Kooperation von Hochschulen und disziplinspezifischen Angeboten beim Management, der Bereitstellung und der langfristigen Sicherung von Forschungsdaten zu intensivieren – etwa durch Vereinbarungen zur Veröffentlichung von Daten in bestimmten fachlichen Repositorien, die Einbindung von externer Expertise in Lehrveranstaltungen (z.B. im Rahmen von Gastvorträgen) oder auch bei der Beratung von Forschenden zu Spezialthemen wie dem Datenschutz.

36 <http://www.qualiservice.org/>.

37 Vgl. Doris Bambej und Marc Rittberger, „Das Forschungsdatenzentrum (FDZ) Bildung des DIPF: Qualitative Daten der empirischen Bildungsforschung im Kontext.“ in *Forschungsinfrastrukturen für die qualitative Sozialforschung*, hrsg. von Denis Huschka et al., 2013, 63-72, zuletzt geprüft am 14.02.2018, https://www.germandataforum.org/dl/downloads/forschungsinfrastrukturen_qualitative_sozialforschung.pdf.

38 <https://www.fdz-bildung.de>. Vgl. Rat für Informationsinfrastrukturen, *Leistung aus Vielfalt: Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland*, 2016, zuletzt geprüft am 14.02.2018, <http://www.rfii.de/?p=1998>.

Literaturverzeichnis

- Akdeniz, Esra und Wolfgang Zenk-Möltgen. *DDI-Lifecycle im Datenarchiv: das Metadatenchema für die Dokumentation in verschiedenen Softwaresystemen*. GESIS Papers, Nr. 2 (2017). <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-50354-9>.
- Arbeitsgruppe Datenschutz und qualitative Sozialforschung Working. *Datenschutzrechtliche Anforderungen bei der Generierung und Archivierung qualitativer Interviewdaten*. Working Paper Series des Rates für Sozial- und Wirtschaftsdaten, Nr. 238 (2014). <http://hdl.handle.net/10419/97181>.
- Article 29 Data Protection Working Party. „Opinion 05/2014 on Anonymisation Techniques.“ 0829/14/EN WP216 (2014). Zuletzt geprüft am 16.02.2018. http://collections.internetmemory.org/haeu/20140424090503/http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.
- Bambey, Doris und Marc Rittberger. „Das Forschungsdatenzentrum (FDZ) Bildung des DIPF: Qualitative Daten der empirischen Bildungsforschung im Kontext.“ In *Forschungsinfrastrukturen für die qualitative Sozialforschung*, herausgegeben von Denis Huschka, Hubert Knoblauch, Claudia Oellers und Heike Solga, 63-72, 2013. Zuletzt geprüft am 14.02.2018. https://www.germandataforum.org/dl/downloads/forschungsinfrastrukturen_qualitative_sozialforschung.pdf.
- Bundesministerium für Bildung und Forschung (BMBF). „Bekanntmachung Richtlinie zur Förderung von Forschung zu ‚Digitalisierung im Bildungsbereich – Grundsatzfragen und Gelingensbedingungen‘.“ Bundesanzeiger AT 26.09.2017 B2 (2017). Zuletzt geprüft am 14.02.2018. <https://www.bmbf.de/foerderungen/bekanntmachung-1420.html>.
- „bwFDM - Koordiniertes Forschungsdatenmanagement in Baden-Württemberg.“ Zuletzt geprüft am 14.02.2018. <https://bwfdm.scc.kit.edu/>.
- Cabrera, Noemi Betancort und Elmar Haake. *Das Qualiservice Metadatenchema. Version 1.0*. Qualiservice Technical Reports, Nr. 13 (2013). <http://nbn-resolving.de/urn:nbn:de:gbv:46-00103310-19>.
- Curdt, Constanze, Volker Hess, Ania Lopez, Benedikt Magrean, Dominik Rudolph und Johanna Vompras. „Herausforderung Forschungsdatenmanagement – Unterstützung der Hochschulen durch eine einrichtungsübergreifende Kooperation in NRW.“ In *E-Science-Tage 2017: Forschungsdaten managen*, herausgegeben von Jonas Kratzke und Vincent Heuveline. Heidelberg: heiBOOKS, 2017. <https://doi.org/10.11588/heibooks.285.377>.

- Deutsche Forschungsgemeinschaft. *Sicherung guter wissenschaftlicher Praxis / Safeguarding Good Scientific Practice. Denkschrift / Memorandum*, 2013. Zuletzt geprüft am 10.03.2018. http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf.
- Deutsche Forschungsgemeinschaft. *Umgang mit Forschungsdaten: DFG-Leitlinien zum Umgang mit Forschungsdaten*, 2015. Zuletzt geprüft am 27.06.2018. http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/.
- Deutsche Gesellschaft für Soziologie und Berufsverband Deutscher Soziologinnen und Soziologen. „Ethik-Kodex der Deutschen Gesellschaft für Soziologie (DGS) und des Berufsverbandes Deutscher Soziologinnen und Soziologen (BDS)“, 2017. Zuletzt geprüft am 10.03.2018. <http://www.soziologie.de/de/die-dgs/ethik/ethik-kodex.html>.
- Eisentraut, Marcus. „Data anonymization.“ In *Data Processing and Documentation: Generating High Quality Research Data in Quantitative Social Science Research* (GESIS Papers), herausgegeben von Sebastian Netscher und Christina Eder. Köln: GESIS - Leibniz-Institut für Sozialwissenschaften, in Vorbereitung.
- Europäische Menschenrechtskonvention (Konvention zum Schutze der Menschenrechte und Grundfreiheiten) vom 04.11.1950. Zuletzt geprüft am 14.02.2018. <https://dejure.org/gesetze/MRK>.
- European Commission. *H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020*. 2016. Zuletzt geprüft am 14.02.2018. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.
- Harzenetter, Karoline. „Metadata documentation at the variable level.“ In *Data Processing and Documentation: Generating High Quality Research Data in Quantitative Social Science Research* (GESIS Papers), herausgegeben von Sebastian Netscher und Christina Eder. Köln: GESIS - Leibniz-Institut für Sozialwissenschaften, in Vorbereitung.
- Huschka, Denis, und Claudia Oellers. „Einführung: Warum qualitative Daten und ihre Sekundäranalyse wichtig sind.“ In *Forschungsinfrastrukturen für die qualitative Sozialforschung*, herausgegeben von Denis Huschka, Hubert Knoblauch, Claudia Oellers und Heike Solga, 9-18. Berlin: Scivero, 2013. Zuletzt geprüft am 14.02.2018. https://www.germandataforum.org/dl/downloads/forschungsinfrastrukturen_qualitative_sozialforschung.pdf.
- Jakobsson, Ulf, Ricarda Braukmann und Malin Lundgren. „Expert tour guide on Data Management. Chapter 1: Plan.“ 2017. Zuletzt geprüft am 14.02.2018. <http://www.cessda.eu/DMGuide>.

- Jensen, Uwe. *Leitlinien zum Management von Forschungsdaten. Sozialwissenschaftliche Umfragedaten*. GESIS Technical Reports, Nr. 7 (2012). <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-320650>.
- Knoblauch, Hubert. „Einigen Anforderungen an Forschungsinfrastrukturen aus der Sicht der qualitativen Forschung“. In *Forschungsinfrastrukturen für die qualitative Sozialforschung*, herausgegeben von Denis Huschka, Hubert Knoblauch, Claudia Oellers und Heike Solga, 27-34. Berlin: Scivero, 2013. Zuletzt geprüft am 14.02.2018. https://www.germandataforum.org/dl/downloads/forschungsinfrastrukturen_qualitative_sozialforschung.pdf.
- Koch, Ute, Esra Akdeniz, Jana Meichsner, Brigitte Hausstein und Karoline Harzenetter. *dalra Metadata Schema Documentation for the Publication and Citation of Social and Economic Data Version 4.0*. GESIS Papers, Nr. 25 (2017). <https://doi.org/10.4232/10.mdsdoc.4.0>.
- Meyermann, Alexia und Maike Porzelt. *Hinweise zur Anonymisierung von qualitativen Daten*. forschungsdaten bildung informiert, Nr. 1 (2014). Zuletzt geprüft am 14.02.2018. https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert-nr-1.pdf.
- Rat für Informationsinfrastrukturen. *Leistung aus Vielfalt: Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland*, 2016. Zuletzt geprüft am 14.02.2018. <http://www.rfii.de/?p=1998>.
- Rat für Sozial- und Wirtschaftsdaten. *Forschungsdatenmanagement in den Sozial-, Verhaltens- und Wirtschaftswissenschaften. Orientierungshilfen für die Beantragung und Begutachtung datengenerierender und datennutzender Forschungsprojekte*. RatSWD Output Series, Nr. 3 (2016). <https://doi.org/10.17620/02671.7>.
- Schaar, Katrin. *Was hat die Wissenschaft beim Datenschutz künftig zu beachten? Allgemeine und spezifische Änderungen beim Datenschutz im Wissenschaftsbereich durch die neue Europäische Datenschutzgrundverordnung*. Working Paper Series des Rates für Sozial- und Wirtschaftsdaten, Nr. 257 (2016). <https://doi.org/10.17620/02671.19>.
- Schwarz, Hannah. „Data consistency.“ In *Data Processing and Documentation: Generating High Quality Research Data in Quantitative Social Science Research* (GESIS Papers), herausgegeben von Sebastian Netscher und Christina Eder. Köln: GESIS - Leibniz-Institut für Sozialwissenschaften, in Vorbereitung.
- Verbund Forschungsdaten Bildung. *Hinweise zur Erstellung eines Methoden- / Feldberichts*. fdbinfo, Nr. 3 (2015). Zuletzt geprüft am 14.02.2018. https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdbinfo_3.pdf.

- Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (DSG) (2016). Zuletzt geprüft am 14.02.2018. <http://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32016R0679&from=DE>.
- Watteler, Oliver. *Erstellung von Methodenberichten für die Archivierung von Forschungsdaten*, 2010. Zuletzt geprüft am 14.02.2018. https://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/Aufbau_Methodenbericht_v1_2010-07.pdf.
- Watteler, Oliver und Katharina E. Kinder-Kurlanda. „Anonymisierung und sicherer Umgang mit Forschungsdaten in der empirischen Sozialforschung.“ *Datenschutz und Datensicherheit: DuD; Recht und Sicherheit in Informationsverarbeitung und Kommunikation* 39, Nr. 8 (2015): 515-19. <https://doi.org/10.1007/s11623-015-0462-0>.
- Wirth, Heike. *Anonymisierung des Mikrozensuspanels im Kontext der Bereitstellung als Scientific-Use-File*, 2006. Zuletzt geprüft am 14.02.2018. <https://www.destatis.de/DE/Methoden/Methodenpapiere/Mikrozensus/Arbeitspapiere/Arbeitspapier11.html>.
- Wirth, Heike. „Die faktische Anonymität von Mikrodaten: Ergebnisse und Konsequenzen eines Forschungsprojektes.“ *ZUMA Nachrichten* 16, Nr. 30 (1992): 7-65. <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-209679>.
- Wissenschaftsrat. *Bericht zur wissenschaftsgeleiteten Bewertung umfangreicher Forschungsinfrastrukturvorhaben für die Nationale Roadmap*, 2017. Zuletzt geprüft am 14.02.2018. <https://www.wissenschaftsrat.de/download/archiv/6410-17.pdf>.
- Zenk-Möltgen, Wolfgang und Norma Habel. *Der GESIS Datenbestandskatalog und sein Metadatenchema. Version 1.8*. GESIS Technical Reports, Nr. 1 (2012). Zuletzt geprüft am 14.02.2018. <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-292372>.

Feldnotizen und Videomitschnitte

Zum Forschungsdatenmanagement qualitativer Daten am Beispiel der ethnologischen Fächer

Matthias Harbeck, Universitätsbibliothek Humboldt-Universität zu Berlin

Sabine Imeri, Universitätsbibliothek Humboldt-Universität zu Berlin

Wjatscheslaw Sterzer, Universitätsbibliothek Humboldt-Universität zu Berlin

Zusammenfassung:

Sollen Daten aus qualitativer (Sozial-)Forschung dauerhaft archiviert und für die Nachnutzung zur Verfügung gestellt werden, dann entstehen erhebliche Probleme: Solche Daten sind kaum standardisiert und liegen in heterogenen Formaten vor, sie sind meist hoch sensibel und benötigen besondere Schutzmechanismen. Der Beitrag thematisiert das Forschungsdatenmanagement qualitativer Daten am Beispiel ethnografischer Forschung aus unterschiedlichen Perspektiven: Er bietet erstens Einblicke in die Positionen von Ethnologinnen und Ethnologen und skizziert fachspezifische Probleme und stellt zweitens vor, welche infrastrukturellen Lösungen für ähnlich gelagerte Schwierigkeiten bereits existieren. Grundlage sind Erhebungen, die im Rahmen des Fachinformationsdienstes Sozial- und Kulturanthropologie an der Universitätsbibliothek der Humboldt-Universität zu Berlin einerseits den aktuellen Stand des Umgangs mit Forschungsdaten und andererseits Bedingungen und Lösungen für die Überführung ethnografischer Materialien in digitale Langzeitarchive eruieren.

Summary:

There are substantial obstacles when speaking of archiving and making reusable qualitative research data from the social sciences. Not only is that kind of data barely standardized and is usually collected in heterogeneous formats, but also due to its highly sensitive nature, adequate access control systems are imperative. The paper sheds light on the research data management of ethnographic research from multiple perspectives: From the lens of ethnographers, the argument first outlines common obstacles in the field and, second, introduces existent digital information infrastructures with viable solutions for similar challenges. The Specialized Information Service Social- und Cultural Anthropology (University Library of Humboldt-Universität zu Berlin) collected comprehensive data to elicit facts about how ethnographers integrate research data management into their daily research practices and what conditions and solutions are deemed necessary to ingest ethnographic material for digital long-term preservation.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S123-141>

Autorenidentifikation: Harbeck, Matthias: GND 1143605314; ORCID <https://orcid.org/0000-0002-7909-5967>; Imeri, Sabine: GND 1153851903; ORCID <https://orcid.org/0000-0002-8844-4014>; Sterzer, Wjatscheslaw: GND 1153852373; ORCID <https://orcid.org/0000-0002-0567-3250>

Schlagwörter: Forschungsdatenmanagement, Fachinformationsdienst, Langzeitarchivierung, Qualitative Sozialforschung, Ethnographie, Sozial- und Kulturanthropologie

1. Forschungsdatenmanagement als eine Aufgabe des FID Sozial- und Kulturanthropologie

Der Umstieg von den Sondersammelgebieten zu den Fachinformationsdiensten (FID) ist vielfach und in manchem zu Recht kritisiert worden. Der Paradigmenwechsel zu einer Projektkultur, die keine Kontinuität gewährleistet und hohen Antragsaufwand erzeugt, wird als Gefährdung und Belastung einer funktionierenden Forschungsinfrastruktur im Bereich der Literatur- und Informationsversorgung gesehen. Gleichwohl sollte ein großer Vorteil der neuen Programmförderung nicht übersehen werden: die Flexibilisierung. Wurde bisher unter Informationsversorgung im Wesentlichen die Versorgung mit Printliteratur verstanden, wird dieser Auftrag in den Fachinformationsdiensten viel breiter und flexibler aufgefasst – „Fachinformation“ meint hier nicht nur klassische Formen der Forschungsliteratur, sondern auch bibliografische Daten, digitale Angebote und die sog. Forschungsdaten, also alle Formen fachlich relevanter Information. Zudem ist jetzt ein flexiblerer Einsatz der Mittel möglich: Es kann nicht nur Literatur und deren Bereitstellung finanziert werden, sondern auch Personal und technische Infrastruktur. Dieser breiter gefasste Versorgungsauftrag hat dazu geführt, dass mehrere FID das Thema „Forschungsdatenmanagement“ in ihre Projekte aufgenommen haben. Weil der Stand der Diskussion über Forschungsdaten in den verschiedenen Disziplinen aber sehr heterogen ist, ist das Thema entsprechend unterschiedlich in die FID-Anträge eingegangen.

1.1. Arbeitsaufträge

Die Antragstellung für den FID Sozial- und Kulturanthropologie – in der Nachfolge des Sondersammelgebietes Volks- und Völkerkunde – erfolgte wesentlich auf der Grundlage von Beratungen über mögliche und benötigte Dienste zwischen der Universitätsbibliothek der Humboldt-Universität zu Berlin als Trägerin der Services und den Fachcommunities. Die Auseinandersetzung mit Forschungsdatenmanagement für die ethnologischen Fächer¹ war eines der Anliegen, die Forschende in diesem Prozess formulierten. Zunächst noch recht unbestimmt, wurden dem Thema einerseits hohes Potential, andererseits komplexe Anforderungen zugeschrieben und darüber hinaus wissenschaftspolitisch eine hohe Dringlichkeit. Denn so wie der Wissenschaftsrat für „maximale Publizität digitaler Forschungsdaten“² eintritt, haben die Forschungsförderer längst begonnen, in ihre Richtlinien die Forderung nach zumindest temporärer Archivierung und Möglichkeiten der sekundären Nutzung von Daten aus öffentlich geförderter Forschung einzuarbeiten.³ Gleichzeitig gibt es kein etabliertes Fachrepositorium für ethnografische Forschungsdaten auf nationaler Ebene.

- 1 Unter dem begrifflichen Dach „Sozial- und Kulturanthropologie“ bzw. „ethnologische Fächer“ sind hier die Fachtraditionen der Völkerkunde/Ethnologie und der Volkskunde/Europäischen Ethnologie im deutschsprachigen Raum subsumiert, deren Institutionen auch unter anderen Bezeichnungen wie Empirische Kulturwissenschaft, Populäre Kulturen oder eben Sozial- und Kulturanthropologie, teils in Kombination, zu finden sind.
- 2 Wissenschaftsrat, *Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020*, Drs. 2359-12, (Berlin: WR, 2012): 53, zuletzt geprüft am 28.11.2017, <https://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>.
- 3 Z.B. wurde im Rahmen von Horizon 2020/EU ab der Förderphase 2017 ein Pilotverfahren so ausgeweitet, dass „das Konzept der offenen Forschungsdaten zum Standard wird.“ Vgl. „Horizont 2020: Aktualisiertes Arbeitsprogramm unterstützt Wettbewerbsfähigkeit durch offene Wissenschaft“, European Commission, zuletzt geprüft am 28.11.2017, http://europa.eu/rapid/press-release_MEMO-16-2604_de.htm. In den Leitlinien der DFG heißt es: „Bereits in die Planung eines Projekts sollten Überlegungen einfließen, ob und welche der aus einem Vorhaben resultierenden Forschungsdaten für andere Forschungskontexte relevant sein können und in welcher Weise diese Forschungsdaten anderen Wissenschaftlerinnen und Wissenschaftlern zur Nachnutzung zur Verfügung gestellt

Zunächst aber musste es darum gehen, die gängige Praxis im Umgang mit den Forschungsdaten genauso festzustellen wie die Vorstellungen und Wünsche von Ethnologinnen und Ethnologen hinsichtlich potenzieller Nutzungsmöglichkeiten. Unklar war überdies, wie ein ideales Forschungsdatenmanagement aussehen und in Forschungsroutrinen integriert werden könnte. Damit waren einige Aufträge an den Fachinformationsdienst bereits im Antrag grob umrissen:

- Erheben, wie und in welchen Formaten Ethnologinnen und Ethnologen Daten produzieren und welche Schwierigkeiten im Umgang mit diesen Daten bestehen (können).
- Eruiieren, wie diese Daten aus üblichen Forschungstools extrahiert und archiviert werden können, idealerweise in einer auf qualitative ethnografische Daten spezialisierten Einrichtung.
- Handlungsempfehlungen oder Best Practice-Beispiele zum Umgang mit und zur Archivierung von Forschungsdaten für die Forschenden erarbeiten.

1.2. Erste Erkenntnisse

Im Zuge der Arbeiten – so viel sei hier vorweggenommen – stellte sich schnell heraus, dass für die Ethnologien zutrifft, was sich auch genereller formulieren lässt: Zwischen dem, was Forschungsförderer einfordern (werden), was in Infrastruktureinrichtungen diskutiert wird und überdies technisch möglich ist, und dem, was methodisch und fachkulturell in vielen Fachdisziplinen üblich ist, besteht eine erhebliche Differenz.⁴ Während es in manchen Disziplinen, wie etwa den Umweltwissenschaften, aber durchaus umfassende Erfahrungen mit Forschungsdatenmanagement und Datenpublikationen gibt, ist das im Bereich der qualitativen (Sozial-)Forschungsdaten – zu denen Daten aus den ethnologischen Fächern überwiegend zählen – kaum der Fall.⁵ Vor allem konkrete Anwendungskenntnisse und vertieftes Wissen sind in den Ethnologien wenig verbreitet, wie wir aus einer Online-Umfrage wissen, die wir im Rahmen des FID Anfang 2016 durchgeführt haben (Abb. 1).⁶ In die Auswertung sind 270 Fragebögen von Forschenden aller Statusgruppen eingeflossen; ein Drittel der Teilnehmenden hat die Umfrage vorzeitig abgebrochen, so dass 181 bis zur letzten Frage ausgefüllte Bögen vorliegen.

werden können.“ „Umgang mit Forschungsdaten: DFG-Leitlinien zum Umgang mit Forschungsdaten“, DFG, zuletzt geprüft am 27.11.2017, http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/index.html.

- 4 Vgl. Ben Kaden, „Dienste für die dissertationsbegleitende Publikation von Forschungsdaten. Eine Vortragsnachlese.“ LIBREAS. *Library Ideas* (Blog), 07. Juni 2017, zuletzt geprüft am 27.11.2017, <https://libreas.wordpress.com/2017/06/07/forschungsdaten-2/#more-4917>.
- 5 Vgl. aus der Perspektive der Sozialwissenschaften die Beiträge in Denis Huschka, Hubert Knoblauch, Claudia Oellers und Heike Solga, Hrsg., *Forschungsinfrastrukturen für die qualitative Sozialforschung* (Berlin: SCIVERO Verlag, 2013), zuletzt geprüft am 13.02.2018, https://www.ratswd.de/dl/downloads/forschungsinfrastrukturen_qualitative_sozialforschung.pdf.
- 6 Sabine Imeri, Ida Danciu [Mitarb.], *Open Data. Forschungsdatenmanagement in den ethnologischen Fächern. Auswertung einer Umfrage des Fachinformationsdienstes Sozial- und Kulturanthropologie an der Universitätsbibliothek der Humboldt-Universität zu Berlin 2016, Teil I: Statistiken* (Version 1.0 / Juni 2017), zuletzt geprüft am 18.06.2018, http://www.evifa.de/cms/fileadmin/uploads/Umfrage_Bericht_Statistiken_1.0_14-06-2017.pdf. Ganz ähnlich haben lediglich zehn Prozent der Teilnehmenden angegeben, für ihr aktuelles Forschungsprojekt einen Datenmanagementplan erstellt zu haben, weitere rund zehn Prozent haben die Absicht, dies zu tun. 48 % haben keinen und rund 20 % haben keine Vorstellung, was ein Datenmanagementplan ist.

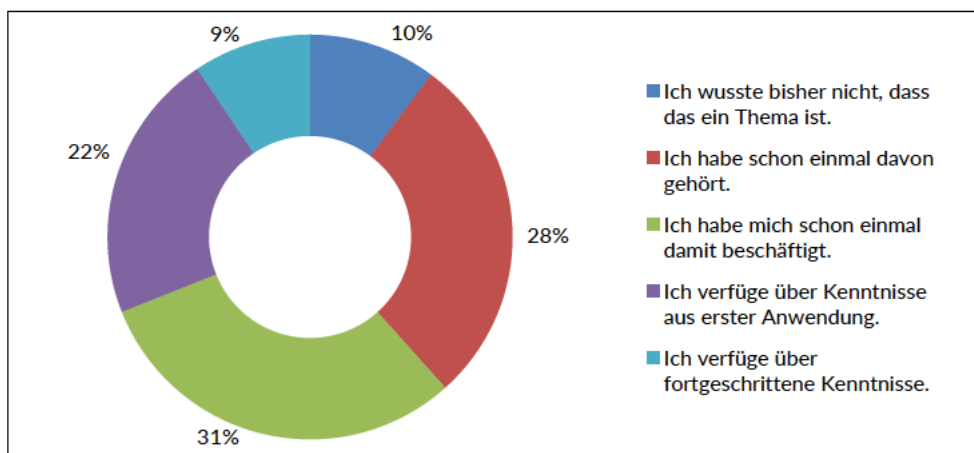


Abb. 1: Wie gut kennen Sie sich im Thema Forschungsdatenmanagement aus? (N=255). Aus Imeri und Danciu, Open Data. Forschungsdatenmanagement in den ethnologischen Fächern: 9.

Dass innerfachliche Diskussionen zum Forschungsdatenmanagement erst noch zu führen sind, zeigt sich z.B. auch in dem Umstand, dass die Ethnologien bisher nicht über fachspezifische, von den Fachgesellschaften getragene Positionspapiere und Empfehlungen zum Umgang mit Forschungsdaten verfügen.⁷ Überdies ist der Umgang mit solchen Daten auch in den informationswissenschaftlichen Debatten deutlich unterrepräsentiert.

Der FID sieht sich deshalb mit zusätzlichen Aufgaben konfrontiert: In Anbetracht der wissenschafts- und förderpolitischen Brisanz bei gleichzeitig relativer Unkenntnis in den ethnologischen Fachcommunities versuchen wir, das Thema vermehrt auf die Agenda zu bringen und zu diskutieren. Wir versuchen auch, gemeinsam mit den relevanten Fachgesellschaften und dem Beirat des FID einen Prozess anzustoßen, an dessen Ende zum einen eine differenzierte, den Eigenheiten der Disziplinen gerecht werdende Positionierung zu den Forderungen der Forschungsförderer und zum anderen adäquate Anforderungen an Datenrepositorien stehen können.

1.3. Ausstattung

Für die Umsetzung wurden zwei Mitarbeitende eingestellt – mit je unterschiedlichen Profilen, um sowohl die fachwissenschaftlich-ethnologischen als auch die technisch-informationswissenschaftlichen Aspekte des Themas abdecken und kommunizieren zu können – und verschiedene Kooperationspartner gesucht. Dazu gehören einerseits Forschungsprojekte, mit deren Unterstützung Routinen im Umgang mit den Daten, „Knackpunkte“ für die Integration von Forschungsdatenmanagement in Forschungsprozesse sowie konkrete Beispiele für übergreifende Probleme erhoben werden. Mit der Hilfe von Infrastruktureinrichtungen und eines Softwareherstellers werden andererseits technische Anforderungen und Routinen getestet sowie etwa vorhandene Lösungen bewertet.

⁷ Das trifft ähnlich auf viele andere Disziplinen zu. Die DFG veröffentlicht auf ihren Webseiten entsprechende Richtlinien. DFG, „Umgang mit Forschungsdaten“.

Erhebungsinstrumente sind vor allem Umfragen und Interviews, ein großer Teil der Ressourcen fließt in Kommunikation und Diskussion sowohl mit Ethnologinnen und Ethnologen als auch mit der Forschungsdatencommunity, z.B. in eigenen thematisch fokussierten Workshops, die beide Communities zusammenbringen und in beide Richtungen wirken sollen.

Die folgenden Ausführungen bündeln entsprechend Zwischenergebnisse aus zwei Arbeitsfeldern: der Arbeit mit Forschenden und den Recherchen in vorhandenen Infrastrukturen. Zunächst aber sollen qualitative Forschung generell und ethnografische Arbeitsweisen im Besonderen in wenigen Stichworten charakterisiert werden.

2. Qualitative Forschung – Qualitative Daten

Ethnografische Forschung erarbeitet ihre Erkenntnisse überwiegend auf der Grundlage qualitativer Methoden und Verfahren. „Qualitative (Sozial-)Forschung“ als Begriff bezeichnet dabei zunächst ganz allgemein die Abgrenzung zur „quantitativen (Sozial-)Forschung“. Während letztere darauf zielt, soziale und gesellschaftliche Phänomene zählend und messend mit Hilfe statistischer Verfahren zu erfassen, also zu quantifizieren, ist es das zentrale Anliegen qualitativer Forschung, Denken und Handeln von Menschen und Gruppen durch Beobachtung nah an alltäglicher sozialer und gesellschaftlicher Praxis zu beschreiben und zu analysieren. Entsprechend ist die zentrale Erhebungsmethode quantitativer Forschungen die standardisierte Befragung, die schriftlich oder mündlich erfolgt und auf der Grundlage von Fragebögen stark strukturierte Daten erzeugt.⁸ Der qualitativen Forschung steht hingegen ein breites, offenes Methodenspektrum zur Verfügung, wie z.B. Formen der Beobachtung, Einzelinterviews oder Gruppendiskussionen. Entsprechend vielfältig sind Formen und Formate von Daten die in den Ethnologien entstehen, das Spektrum reicht von Beobachtungsprotokollen über Transkripte, Fotografien und Videoaufzeichnungen bis zu Daten aus den sozialen Medien (Abb. 2).

2.1. Ethnografische Forschungspraxis

Ethnologinnen und Ethnologen arbeiten vor allem mit solchen qualitativen Methoden, sie betreiben in vielen Forschungsszenarien *teilnehmende Beobachtung*, eine Form der Datengewinnung, die auf „Begegnung, Interaktion und der sozialen Teilnahme am Alltagsleben unterschiedlicher Menschen“ beruht.⁹ Umfassender bezeichnet *Feldforschung* den gesamten Forschungsprozess, in dem teilnehmende Beobachtung mit Interviewmethoden, Kartierungen und anderen Formen der Dokumentation, aber auch Archivstudien kombiniert werden. Mit der Feldforschung verbunden ist ein dezidierter Methodenpluralismus, der oft diverses, sich gegenseitig kommentierendes Datenmaterial generiert. Denn ganz gleich, wo auf der Welt Ethnologinnen und Ethnologen forschen, und ganz gleich, ob sie sich für Migration und Mobilität interessieren, für Formen und Netzwerke von Religiosität, für Gesundheitsverhalten, Geschlechterverhältnisse und Sexualität oder für Phänomene der Medialisierung und

8 Vgl. z.B. Jürgen Raithel, *Quantitative Forschung: Ein Praxiskurs*. 2., durchgesehene Auflage (Wiesbaden: VS Verlag für Sozialwissenschaften, 2008), <https://doi.org/10.1007/978-3-531-91148-9>.

9 Michi Knecht, „Nach Writing Culture, mit Actor-Network: Ethnographie/Praxeographie im Feld der Wissenschafts-, Medizin- und Technikanthropologie,“ in *Europäisch-ethnologisches Forschen: Neue Methoden und Konzepte*, hrsg. Sabine Hess, Johannes Moser und Maria Schwertl (Berlin: Reimer, 2013), 79–106.

Technisierung von Alltag¹⁰, benötigt wird Material, das geeignet ist, die Komplexität sozialer, soziotechnischer und sozio-natürlicher Ordnungen angemessen zu beschreiben.¹¹

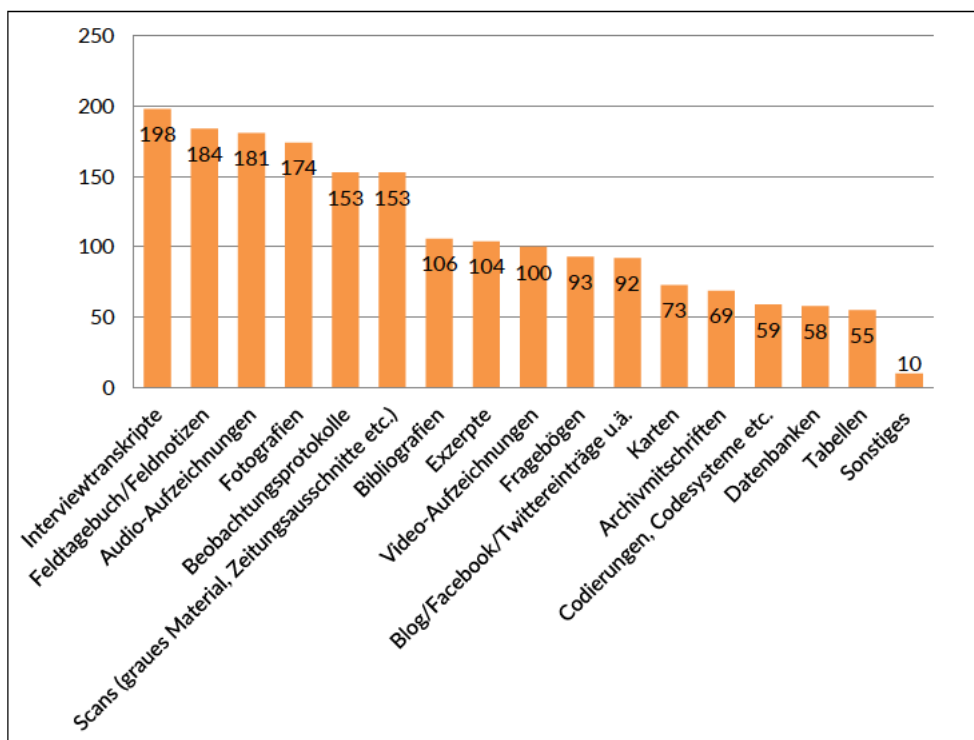


Abb. 2: Was sind in Ihrer (aktuellen oder abgeschlossenen) Forschung Forschungsdaten? (N=251). Mehrfachantworten waren möglich. Aus Imeri und Danciu, Open Data. Forschungsdatenmanagement in den ethnologischen Fächern: 10.

Feldforschungen werden als nicht replizierbare, weil situations- und beobachterabhängige, offene und flexible, oft unvorhersehbar verlaufende Prozesse konzipiert. Feldforschung bedeutet in der Regel, dass Forschende die Lebensräume der Protagonistinnen und Protagonisten ihrer Forschung persönlich aufsuchen. Ein „gründliches, aufwändiges Sich-Einlassen auf *real-world-situations*“¹² ist dabei konstitutiv. Entsprechend entsteht ein großer Teil ethnografischer Daten in einem Prozess, in dem wechselseitig intensive, teils langfristige Beziehungen aufgebaut werden. In jedem Fall sind Forschende in einem viel stärkeren Maße auf die Kooperation und Zustimmung der Akteurinnen und

10 Zu den Themenfeldern, die derzeit in den ethnologischen Fächern schwerpunktmäßig bearbeitet werden vgl. die Kommissionen und Arbeitsgruppen der Fachgesellschaften: „Kommissionen,“ Deutsche Gesellschaft für Volkskunde, zuletzt geprüft am 27.11.2017, <http://www.d-g-v.org/kommissionen>, und „Arbeits- und Regionalgruppen der DGSKA,“ Deutsche Gesellschaft für Sozial- und Kulturanthropologie, zuletzt geprüft am 19.06.2018, <https://www.dgska.de/arbeitsgruppen-2/>.

11 Michi Knecht, „Nach Writing Culture“.

12 Ebd., 91.

Akteure im Feld angewiesen, als das bei großflächigen Datenerhebungen mit Behördenunterstützung wie beim *Nationalen Bildungspanel* oder beim *Zensus* der Fall ist. Feldforschungen sind oft Einzelforschungen, die wesentlich davon leben, dass ein Feldzugang überhaupt erarbeitet werden kann. Forschungsbeziehungen in der Ethnografie sind deshalb „fragile Gebilde“¹³ und Vertrauensverhältnisse, in denen in mehrfacher Hinsicht sensible Daten produziert werden: Forschende erhalten intensive Einblicke in die Alltage von Personen und Gruppen, erfahren Details, z.B. über religiöse Praktiken, sexuelle Orientierungen oder politische Haltungen. Mitunter können ihr Wissen und ihre Daten zu Risiken für die beforschten Akteurinnen und Akteure werden, in eskalierenden Konflikten, in Krisenregionen oder Milieus am Rand der Legalität. Sensibel sind ethnografische Daten aber auch mit Blick auf die Forschenden selbst, weil z.B. die Emotionen der Forschenden eine wichtige Rolle im Erkenntnisprozess spielen können.¹⁴ Und das bedeutet eben auch, dass Forschende selbst in ihren Daten als Personen erkennbar werden.

Nicht zuletzt deshalb ist die Auseinandersetzung mit forschungsethischen Fragen ein wichtiger Teil der Fachdebatten in den Ethnologien¹⁵, gefolgt von einer hohen Sensibilität für die Anforderungen verantwortungsvollen Forschens, auch und gerade mit Blick auf den Umgang mit Daten und deren Verwendung nach dem Abschluss einer Forschung. Auch hier müssen potenzielle Folgen bedacht und abgewogen werden.

2.2. Forschungsdaten in den Ethnologien: Probleme und Anforderungen

In unseren Umfrageergebnissen wird deutlich, dass Ethnologinnen und Ethnologen insgesamt dem Datenmanagement zwar häufig kritisch, aber nicht grundsätzlich ablehnend gegenüber stehen.¹⁶ Das Spektrum der Einschätzungen vor allem zu Langzeitarchivierung und Nachnutzung von Forschungsdaten durch Dritte ist gleichwohl enorm breit: Stark ablehnende Haltungen begegnen hier ebenso wie Nachfragen nach konkreten infrastrukturellen Angeboten zur Datenarchivierung, Einwände haben zunächst eine forschungsökonomische Dimension, beziehen sich also – wie das auch aus anderen Umfragen bekannt ist¹⁷ – auf den zu erwartenden Arbeitsaufwand oder die Finanzierung der Datenaufbereitung. Es werden darüber hinaus aber fachspezifische Problemlagen erkennbar, die aus den methodischen Besonderheiten ethnografischer Feldforschung resultieren. Die in den folgenden

13 Georg Breidenstein et al., *Ethnografie: Die Praxis der Feldforschung*, 2. Auflage (Konstanz: UVK Verlagsgesellschaft mbH, 2015), 62.

14 Thomas Stodulka, „Feldforschung als Begegnung – Zur pragmatischen Dimension ethnographischer Daten,“ *Sociologie* 64, Nr. 2 (2014): 179–206, <https://doi.org/10.3790/soc.64.2.179>.

15 Dies ist auch ein Ergebnis einer seit mehreren Jahrzehnten anhaltenden und unabgeschlossenen Auseinandersetzung mit den Bedingungen, den historischen und politischen Kontexten und Rahmungen ethnografischer Wissensproduktion. Dazu im Überblick z.B. Hansjörg Dilger, „Ethik und Ethnologie: Ethikkommissionen, ethnographisches Arbeiten und Epistemologie – nicht nur in der Medizinethnologie,“ *Medizinethnologie: Körper, Gesundheit und Heilung in einer globalisierten Welt* (Blog), 16. September 2015, zuletzt geprüft am 27.11.2017, <http://www.medizinethnologie.net/ethik-und-ethnologie/>. Manche Fachgesellschaften haben eigene Erklärungen und Richtlinien zur Forschungsethik erarbeitet, vgl. z.B. „Ethikerklärung der DGSKA,“ Deutsche Gesellschaft für Sozial- und Kulturanthropologie, zuletzt geprüft am 19.06.2018, <https://www.dgska.de/dgska/ethik/>, oder die „Principles of Professional Responsibility“ der American Anthropological Association (2012), zuletzt geprüft am 27.11.2017, <http://ethics.americananthro.org/category/statement/>.

16 Sabine Imeri, Ida Danciu [Mitarb.], *Open Data. Forschungsdatenmanagement in den ethnologischen Fächern*, 19, 22 und 24.

17 Vgl. z.B. Patrick J. Droß, *Kurzstudie: Anforderungen an die Archivierung sozial- und wirtschaftswissenschaftlicher Forschungsdaten* (Berlin: WZB, 2015), zuletzt geprüft am 28.11.2017, https://sowidatanet.de/images/pdfs/Meldungen/Kurzstudie_Qualitative_Interviews.pdf.

Passagen verwendeten Zitate sind Freitext-Kommentaren aus unserer Umfrage entnommen. Sie geben exemplarisch Einschätzungen wieder, die sich ähnlich in zahlreichen anderen Kommentaren finden.

2.2.1. Erhebungskontexte dokumentieren

„Die Besonderheit der Feldforschung ist eben genau die Intersubjektivität der Forschungssituation, die sich anhand von Tonaufnahmen oder Aufzeichnungen weder ‚überprüfen‘ noch rekonstruieren lässt.“

„Ich finde es schwierig, ethnografische Daten von anderen neu analysieren zu lassen, da jegliche Kontextfaktoren, das situationale Setting und vor allem die eigene Erfahrung nicht in einer Datenbank abgelegt werden können.“

Ob und wie Feldforschungsbeziehungen und die komplexen Kontexte der Datenerhebung so dokumentiert werden können, dass sie für Dritte – vielleicht sogar fachfremde Forschende – verständlich und damit überhaupt sinnvoll nutzbar werden, ist eine offene Frage: „Datasets don’t speak for themselves.“¹⁸ Diskutiert werden müssen demnach fachliche Standards für eine angemessene (forschungsbegleitende) Kontextualisierung ethnografischer Daten, die methodische Zugänge und Feldsituationen möglichst transparent und nachvollziehbar machen. Die Anreicherung mit Metadaten allein wird hier kaum genügen, überdies fehlen derzeit spezifische Metadatenstandards und einheitliche Fach-Thesauri. Zu berücksichtigen wären hier z.B. Fragen des Datenschutzes etwa mit Blick auf geografische Identifizierbarkeit.

2.2.2. Daten für die Archivierung auswählen

„Feldtagebücher können sehr intime Quellen bzw. ein geschützter Raum sein, die eben nicht zur Veröffentlichung gedacht sind, sondern Zweifel, Emotionen, Irrwege etc. beinhalten sollen und müssen.“

„Das Schreiben von Forschungsprotokollen mit der Perspektive auf ‚volle Transparenz‘ ändert die Offenheit und Experimentierfreude der Forschenden.“

Selbstschutz – der durchaus ein legitimes Anliegen ist – oder Vorbehalte gegenüber Überprüfung und Kontrolle sind häufige Bezugspunkte in den Kommentaren. Abbildung 3 zeigt, dass die Bereitschaft, Daten in ein Datenarchiv zu überführen, abnimmt, je mehr das Material tatsächlich oder vermeintlich solches ist, in dem die oder der Forschende als Person erkennbar wird. Besonders mit Blick auf das Feldtagebuch bzw. Feldnotizen, die als hybrides Material in der Regel nicht nur „Rohdaten“ aus Beobachtungen, sondern bereits Interpretationen, aber auch Persönliches, Befindlichkeiten, Aversionen enthalten können,¹⁹ ist die Skepsis besonders groß.

Genereller geht es hierbei aber um die Auswahl von zur Nachnutzung geeigneten Daten und deren Aufbereitung mit vertretbarem Aufwand. Welches Material bei welchem Stand der Aufbereitung es

18 Rena Lederman, „Archiving Fieldnotes? Placing ‘Anthropological Records’ Among Plural Digital Worlds,“ in *eFieldnotes: The Makings of Anthropology in the Digital World*, hrsg. Roger Sanjek und Susan W. Tratner (Philadelphia: University of Pennsylvania Press, 2016), 261.

19 Robert M. Emerson, Rachel I. Fretz und Linda L. Shaw, *Writing Ethnographic Fieldnotes*, 2. Auflage (Chicago, London: University of Chicago Press: 2011).

„wert“ ist, dauerhaft archiviert zu werden, oder welches Material geeignet ist für welche Nutzungsszenarien – dafür müssen fachspezifische Kriterien erst noch entwickelt werden. Feldtagebücher werden möglicherweise nicht oder nur unter sehr spezifischen Bedingungen dazu gehören.²⁰

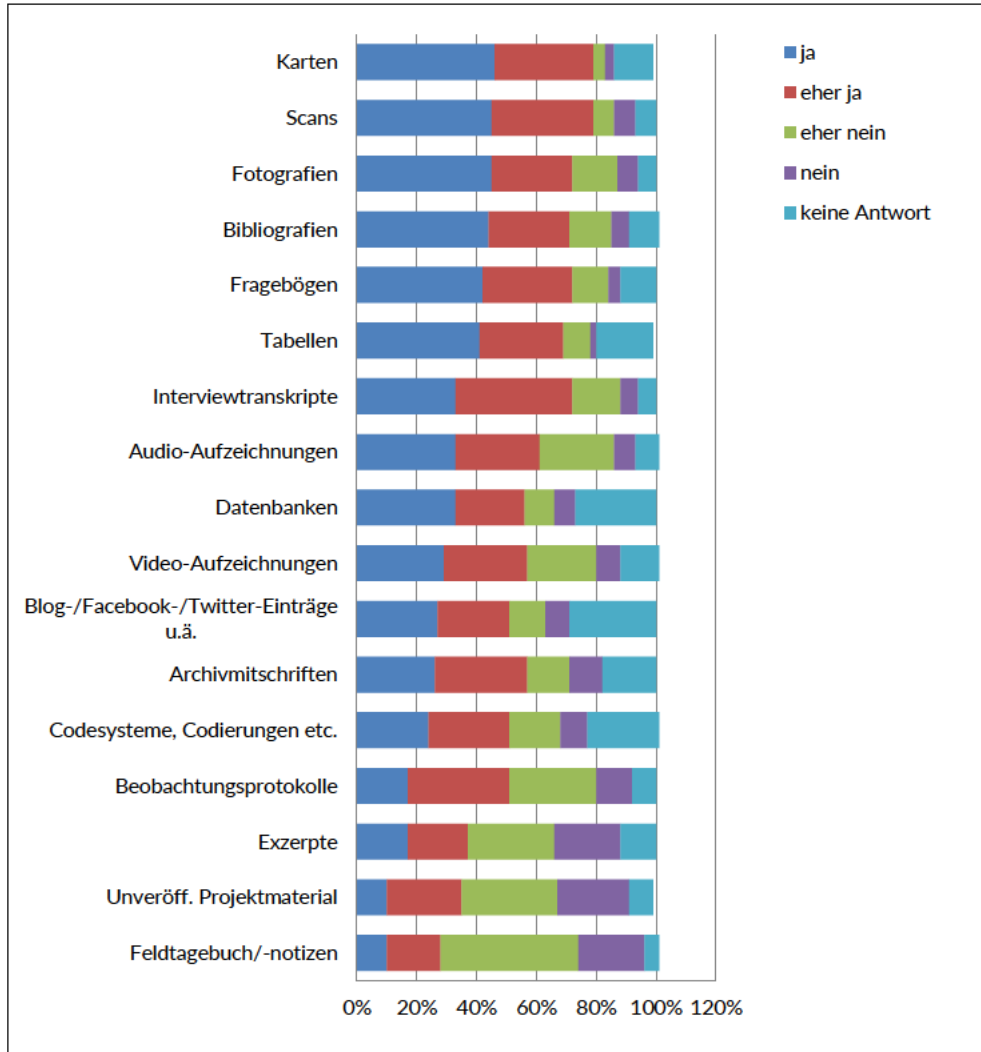


Abb. 3: Würden Sie die folgenden Datentypen in einem Datenarchiv ablegen? (N=181). Aus Imeri und Danciu, Open Data. Forschungsdatenmanagement in den ethnologischen Fächern: 27.

²⁰ Vgl. Rena Lederman, „Archiving Fieldnotes?“, 2016.

2.2.3. Rechtliche Fragen und Forschungsethik

„In vielen ethnologischen Forschungskontexten bedarf es einer Vertrauensbildung [...]. Wenn Transkripte aus solchen Kontexten danach frei zugänglich gemacht werden, unterläuft das systematisch diese Basis.“

„Aus der Fachgeschichte sind Fälle bekannt, in denen Ethnograf/inn/en Forschungsdaten zurückgehalten und auch vernichtet haben, um sie dem Zugriff anderer, auch dem staatlicher Stellen, zu entziehen und damit das Vertrauen, das Forschungspartner/innen in die Forschenden gesetzt haben, zu schützen.“

„Viele Daten sind nicht anonymisierbar, ohne ihren Kontext total zu verwischen – und dann werden sie uninterpretierbar.“

Das mit Blick auf Langzeitarchivierung und Data-Sharing wichtigste und kontroverseste Thema ist die Wahrung der Vertraulichkeit in Verbindung mit Fragen von Daten- und Persönlichkeitsschutz sowie Schnittfelder zu forschungsethischen Fragen.²¹ Strittig ist zum Beispiel, ob und wie in offenen Feldsituationen Einwilligungserklärungen eingeholt werden können, die dann auch auf Szenarien zukünftiger Nutzung der Daten durch unbekannte Dritte ausgedehnt werden müssen. Antworten auf die Frage, wie sich komplexes ethnografisches Material so anonymisieren lässt, dass gleichzeitig die Interpretierbarkeit erhalten bleibt, wird hier eine zentrale Rolle zukommen.²² Zumal sich Konzepte von „Privatheit“ oder „Sensibilität“ von Daten dynamisch entwickeln und bei Datenübergabe an ein Archiv kaum abzusehen ist, auf welchen Wegen offene oder unvollständig anonymisierte Daten künftig Schaden anrichten können.²³ Auch werden vereinzelt Forderungen nach regelrechtem Quellenschutz erhoben, vor allem dann, wenn Ethnografinnen und Ethnografen in Konfliktfeldern wie z. B. Land- und Ressourcenkonflikten oder politischen Protestbewegungen forschen.

Bestenfalls andiskutiert sind Fragen des Urheberrechts. Ein Grund dafür ist, dass die Ausgestaltung des Wissenschaftsurheberrechts generell im Fluss ist.²⁴ Darüber hinaus – und auch das berührt fachspezifische methodologische und vor allem forschungsethische Fragen – werden Daten in den ethnologischen Fächern als von Forschenden und Beforschten ko-produziert angesehen.²⁵ Wenig thematisiert sind auch cultural property rights, die etwa im Diskurs um Bestimmung, Erhalt und

21 Hella von Unger, Hansjörg Dilger und Michael Schönhuth, „Ethikbegutachtung in der sozial- und kulturwissenschaftlichen Forschung? Ein Debattenbeitrag aus soziologischer und ethnologischer Sicht,“ *Forum: Qualitative Sozialforschung / Forum: Qualitative Social Research* 17, Nr. 3 (2016), Art. 20, <http://nbn-resolving.de/urn:nbn:de:0114-fqs1603203>.

22 Vgl. mit Blick auf qualitative Interviews Tobias Gebel et al., „Verboten ist, was nicht ausdrücklich erlaubt ist: Datenschutz in qualitativen Interviews,“ *Forum: Qualitative Sozialforschung / Forum: Qualitative Social Research* 16, Nr. 2 (2015), Art. 27, <http://nbn-resolving.de/urn:nbn:de:0114-fqs1502279>.

23 Vgl. Lisa Cliggett, „Preservation, Sharing and Technological Challenges of Longitudinal Research in the Digital Age,“ in *eFieldnotes: The Makings of Anthropology in the Digital World*, hrsg. Roger Sanjek und Susan W. Tratner (Philadelphia: University of Pennsylvania Press, 2016), 245.

24 Vgl. z.B. Alexander Peukert und Marcus Sonnenberg, „Das Urheberrecht und der Wandel des wissenschaftlichen Kommunikationssystems,“ in *Wissenschaftliches Publizieren. Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung*, hrsg. Peter Weingart und Niels Taubert (Berlin, Boston: De Gruyter Akademie Forschung, 2016), 211–242.

25 Vgl. grundlegend Johannes Fabian, *Time and the Other. How Anthropology Makes Its Object* (New York: Columbia University Press, 1983).

Vermarktung von Kulturerbe zentral sind und im Zusammenhang mit ethnografischen Forschungsdaten und ihrer Verfügbarmachung eine Rolle spielen können.²⁶

2.3. Anforderungen an Datenarchive

Trotz dieser vielen Unklarheiten, dem großen Diskussions- und Regelungsbedarf und der immer wieder geäußerten Skepsis zeigt sich in unseren Umfrageergebnissen gleichwohl eine überwiegend positive Grundhaltung gegenüber Datenarchivierung und -nachnutzung genauso wie die Bereitschaft, sich weiter mit dem Thema zu befassen: Zwei Drittel der von uns Befragten können sich vorstellen, Daten anderer sekundär für die eigene Forschung zu nutzen und ebenso viele sind *unter bestimmten Bedingungen* bereit, künftig Daten in Repositorien zugänglich zu machen. Die Bedingungen für die Zugänglichmachung eigener Daten sind hier jedoch entscheidend. Wie im Grunde überwältigend wichtig Fragen nach Persönlichkeits- und Datenschutz sind, zeigt die Abbildung 4.

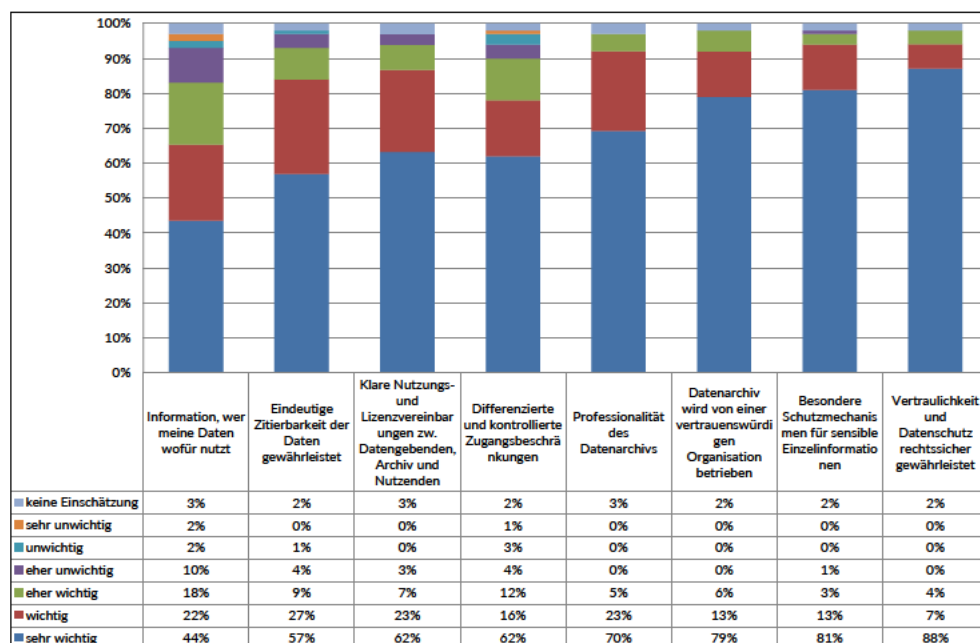


Abb. 4: Wenn Sie Daten in einem Datenarchiv archivieren würden, wie wichtig wären Ihnen folgende Faktoren? (N=181). Aus Imeri und Danciu, *Open Data. Forschungsdatenmanagement in den ethnologischen Fächern*: 25.

Ob Ethnologinnen und Ethnologen ihre Daten in ein Repositorium geben werden, steht und fällt also mit den Antworten, die hier gegeben werden, und den Lösungen, die Datenarchive anbieten

²⁶ Wie z.B. bei der Archivierung von Sprachdaten im Datenarchiv „DOBES. Documentation of Endangered Languages“. Vgl. Thomas Widlok, „The Archive Strikes Back: Effects of Online Digital Language Archiving on Research Relations and Property Rights,“ in *Oral Literature in the Digital Age: Archiving Orality and Connecting with Communities*, hrsg. Mark Turin, Claire Wheeler und Eleanor Wilkinson (Cambridge: Open Book Publishers, 2013), 3–19, zuletzt geprüft am 28.11.2017, <https://www.openbookpublishers.com/reader.php/186?186>.

können. Denn es wird auch befürchtet, dass es sonst schwieriger werden könnte, überhaupt noch Gesprächspartnerinnen und -partner zu finden.

3. Infrastrukturelle Lösungen – Was gibt es?

Während die GESIS – Leibniz-Institut für Sozialwissenschaften bereits seit mehreren Jahrzehnten Daten aus quantitativer Sozialforschung archiviert und zur Sekundärnutzung bereitstellt, stehen für Daten aus qualitativen – und damit auch aus ethnografischen – Forschungen in Deutschland kaum Möglichkeiten der dauerhaften Archivierung zur Verfügung. Deshalb ist es ein wichtiger Teil der Arbeit des FID zu eruieren, welche technischen Lösungen für ähnlich komplexe Probleme in bereits arbeitenden Datenarchiven gefunden worden sind. Der FID zielt dabei nicht auf die Etablierung einer eigenen Infrastruktur, vielmehr geht es darum, in einem ersten Schritt infrastrukturelle Anforderungen aus vorhandenen Workflows und Empfehlungen abzuleiten, um zur Konzeptualisierung fachspezifischer Lösungsansätze beizutragen.

Im Mittelpunkt unserer Beschäftigung mit Infrastrukturen und deren Lösungsansätzen für generelle Probleme stehen Erfahrungswerte und Know-how in den Bereichen Technik, Datenschutz und Urheberrecht sowie Langzeitarchivierung. Von Interesse ist zudem, wie je fachspezifische Probleme mit Blick auf die Kuration heterogener Datentypen organisatorisch und technisch gelöst werden. Mit folgenden Datenarchiven sind wir vor allem wegen ihrer fachlichen oder methodischen Nähe vermehrt im Austausch:

- Forschungsdatenzentrum (FDZ) Bildung am Deutschen Institut für Internationale Pädagogische Forschung (DIPF), Frankfurt am Main
- Datenrepositorium datorium bei der GESIS – Leibniz-Institut für Sozialwissenschaften, Köln
- Forschungsdatenzentrum Ruhr am RWI – Leibniz-Institut für Wirtschaftsforschung, Essen
- IANUS – Forschungsdatenzentrum Archäologie & Altertumswissenschaften, Deutsches Archäologisches Institut Berlin
- Qualiservice, Universität Bremen

3.1. Technische Basis

In der Regel zeigen sich die Betreiber der Datenarchive skeptisch hinsichtlich der funktionalen Kapazitäten vorhandener Open-Source Archivlösungen wie z.B. DSpace. Sie sehen Probleme, etwa bei der Implementierung eines gestaffelten Rechtemanagements, bei der Herausforderung, einzelne Dateien innerhalb von Projekten mit unterschiedlichen Beschreibungen zu versehen, oder berichten über unzureichende Kuratierungsfunktionen für heterogene Datenformate. Diese Schwierigkeiten lassen sich jedoch nicht generalisieren, sie sind vielmehr den fachspezifischen Anforderungen an die Infrastrukturen geschuldet. So etwa, wenn eine eigens entwickelte Oberfläche unabdingbar ist, um die komplexe Differenzierung von Zugriffsrechten qualitativer Forschungsdaten (DIPF) bzw. die Balance zwischen automatisierter und manueller Kuration²⁷ (IANUS) technisch adäquat umzusetzen.

27 Um große Mengen an Forschungsdaten effizient und ressourcenschonend ins Archiv zu überführen, ist Automatisierung wünschenswert. Jedoch kann es Fälle geben, in denen das System bestimmte Datentypen oder Dateiformate nicht erkennt und somit manuelle Eingriffe notwendig werden.

Während einige Infrastruktureinrichtungen die Open-Source-Software ergänzende Entwicklungen vornehmen, setzen andere deshalb ausschließlich auf Eigenentwicklungen der technischen Basis. Das heißt jedoch nicht, dass gänzlich auf Open-Source-Software verzichtet wurde, vielmehr sind einzelne Open-Source-Tools entsprechend erweitert und an den selbst entwickelten Kern angepasst worden. Beide Wege deuten aber auf einen generellen Bedarf an Eigenentwicklung beim Aufbau von Datenarchiven und auf die leider noch nicht ausreichenden Angebote vorgefertigter Lösungen – hier vor allem DSpace – hin.

3.2. Zwischen Datenschutz und Analysepotenzial

Die Anonymisierung von personenbezogenen Forschungsdaten ist die zentrale Methode zur Einhaltung des Datenschutzes aller auf die Sozial-, Wirtschafts- bzw. Bildungsforschung spezialisierten Datenarchive. Dabei werden personenbezogene Daten so unkenntlich gemacht, dass eine Identifizierung nicht oder nur mit erheblichem Aufwand möglich ist. Prinzipiell wird zwischen direkten und indirekten Identifikatoren unterschieden. Direkte Identifikatoren sind jene, die explizit auf die Personen verweisen, wohingegen indirekte Identifikatoren²⁸ Rückschlüsse durch Sekundärinformationen erlauben, die zur Identifizierung der Person führen können. Dabei entsteht immer ein Konflikt, weil die Anonymisierung das analytische Potential der Daten teils erheblich einschränken kann. In den von uns konsultierten Archiven sind die Datengebenden in erster Instanz für die Anonymisierung ihrer Forschungsdaten zuständig. Bei Übergabe werden die Daten seitens der Datenarchive auf Anonymisierungslücken überprüft. Vor allem bei den indirekten Identifikatoren – so die Berichte aus den Archiven – besteht häufig Korrekturbedarf, der im Kontakt mit den Datengebenden behoben wird. Die Archive erfüllen damit nicht nur die rechtlichen Datenschutzaufgaben für die Nachnutzung, sondern leisten auch einen Beitrag zur eigenen Vertrauenswürdigkeit.

Die Balance zwischen adäquatem Daten- und Persönlichkeitsschutz und Nachnutzungspotential zu halten, kann – abhängig vom Datentyp – sehr kompliziert sein. Das FDZ Ruhr am RWI z.B. bietet immer zwei Versionen eines Datensets an, wobei eines stärker anonymisiert ist als das andere. Die stark anonymisierte Version unterliegt weniger strikten Zugriffsrechten. Reichen die vorhandenen Variablen für die geplante Nutzung nicht aus, kann die leicht anonymisierte Version genutzt werden – allerdings nur via On-Site-Zugang: Im sog. *Safe Room* werden hochsensible personenbezogene Daten unter kontrollierten Bedingungen zur Verfügung gestellt. Forschende können Daten nur an Gastarbeitsplätzen vor Ort einsehen, nur bedingt oder gar nicht auf einer virtuellen Maschine bearbeiten, nicht speichern, verschicken oder drucken. Ähnliche Lösungen bieten das Datenrepositorium datorium und Qualiservice an.

Im DIPF – wie bei Qualiservice – wird die sog. *faktische Anonymisierung* angewandt, eine Methode zur Umformulierung von indirekten Identifikatoren, bei der die Merkmale einer Person mit sinnge rechten Pseudonymen so umschrieben werden, dass das analytische Potential der Daten so weit wie

28 Indirekte Identifikatoren beinhalten Informationen, die zur Identifizierung einer Person führen können, z.B. ‚Rektor der örtlichen Grundschule‘ oder ‚Autobauer aus Niedersachsen‘. Vgl. Alexia Meyer mann und Maïke Porzelt, ‚Hinweise zur Anonymisierung von qualitativen Daten,‘ *forschungsdaten bildung informiert*, Nr. 1 (2014): 6, zuletzt geprüft am 19.06.2018, <https://www.forschungsdaten-bildung.de/anonymisierung>.

möglich erhalten bleibt.²⁹ Weil diese Form der Anonymisierung sehr stark vom Fokus des jeweiligen Datensets abhängt und individuell bewertet werden muss, gibt es jedoch kein allgemeingültiges Verfahren. Video- und Audiodaten werden hingegen nicht anonymisiert, Gesichter werden nicht verpixelt und Stimmen nicht verzerrt. Auch hier soll das analytische Potential des ohnehin seltenen Materials erhalten werden, z.B. wenn Gesichtsausdrücke für die Analyse einer Unterrichtssituation ausschlaggebend sind. Auch bei nicht-personenbezogenen Forschungsdaten kann es vorkommen, dass anonymisiert werden muss. Z.B. sind Fälle denkbar, in denen mit Blick auf Kunst- und Kulturgüterschutz Orte archäologischer Fundstellen unkenntlich gemacht werden müssen.

3.3. Urheberrecht: Verschiedene Wege

Wie oben bereits angedeutet, werden die Fragen des Wissenschaftsurheberrechts durchaus kontrovers diskutiert. In Frage steht etwa, ob bestimmte Forschungsdaten Schöpfungshöhe³⁰ überhaupt aufweisen können, was ggf. im Einzelfall geprüft werden muss.³¹ Die Datenarchive finden je eigene Wege, mit Urheberrechtsfragen umzugehen: Die FDZs DIPF und IANUS setzen generell Urheberrechte für Forschungsdaten voraus. Ausschlaggebend ist hier, dass zur Vollständigkeit eines Datensets nicht nur Rohdaten, sondern auch Supplemente gehören, die über Forschungsdesigns – Fragestellungen, Methoden, Vorbereitung etc. – Auskunft geben können. Diese Gesamtheit wird als Ergebnis geistiger Arbeit angesehen, die deshalb Schöpfungshöhe aufweist. Während datorium Forschungsdaten unter der Annahme archiviert, dass Urheberrechte potenziell mindestens für Teile von Datensets geltend gemacht werden könnten, hat das FDZ Ruhr die Entscheidung getroffen, gänzlich auf die Anerkennung von Schöpfungshöhe und Urheberrechten an Forschungsdaten zu verzichten. Wichtiger Grund für diese Entscheidung ist, dass das FDZ Ruhr am RWI als institutionelles Repositorium ausschließlich Daten archiviert, die im RWI selbst produziert werden. Bei der Übergabe werden sämtliche Rechte übertragen, die Namen der einzelnen Forschenden werden jedoch in den Metadaten verzeichnet und somit Zitierbarkeit gewährleistet. Entsprechend werden in allen von uns kontaktierten Datenarchiven bei Übermittlung der Forschungsdaten Datengeberverträge mit den Forschenden geschlossen, um die Rechte und Pflichten für alle Beteiligten zu regeln.

3.4. Langzeitarchivierung

Für die Langzeitarchivierung von Forschungsdaten gibt es keine klaren und allgemeingültigen Richtlinien bezüglich des Zeitraumes und des Vorgehens. Seit 1998 sieht etwa die DFG für die Sicherung und Aufbewahrung von Primärdaten im Rahmen ihrer Leitlinien zur guten wissenschaftlichen Praxis zehn Jahre vor.³² Das Kompetenznetzwerk Langzeitarchivierung nestor formuliert, es ginge bei den Fristen „explizit nicht um einen begrenzten, fest definierten Zeitraum, sondern vielmehr sollte die

29 Ebd., 7; Susanne Kretzer, *Arbeitspapier zur Konzeptentwicklung der Anonymisierung / Pseudonymisierung in Qualiservice* (Bremen: Universität Bremen, 2013), zuletzt geprüft am 28.11.2017, http://www.qualiservice.org/fileadmin/templates/qualiservice/Anonymisierungskonzept_Arbeitspapier.pdf.

30 Schöpfungshöhe meint, dass ein Werk eine persönliche geistige Schöpfung ist und ein Mindestmaß an Originalität aufweist. Urheberrecht setzt Schöpfungshöhe voraus.

31 „Urheberrecht,“ *Forschungsdaten.org*, zuletzt geprüft am 28.11.2017, <http://www.forschungsdaten.org/index.php/Urheberrecht>.

32 DFG, *Sicherung guter wissenschaftlicher Praxis: Denkschrift*, ergänzte Auflage (Weinheim: WILEY-VCH Verlag GmbH & Co. KGaA, 2013), 21, zuletzt geprüft am 28.11.2017, www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf. Die DFG hatte mit diesen Richtlinien allerdings eher wissenschaftliches Fehlverhalten im Blick, also Täuschung oder Plagiat.

Diskussion von dem Anspruch geleitet werden, Forschungsdaten ‚bis auf Weiteres‘ zu archivieren.“³³ „Bis auf Weiteres“ suggeriert, dass Forschungsdaten auf unbestimmte Zeit archiviert werden sollen und eröffnet eine gewisse Flexibilität in der Definition des Zeitraums, den Archive und Repositorien unseren Recherchen zufolge selbst festlegen. Mehr als die Hälfte der Ethnologinnen und Ethnologen, die an unserer Umfrage teilgenommen haben, hat sich z.B. für eine potenziell unbegrenzte Archivierung ausgesprochen. Gleichzeitig plädieren Forschende aber auch für eine Datenauswahl und die Bemessung von Archivierungsfristen an der Qualität der Daten, schließlich gelte es, große unbrauchbare Datenmengen zu verhindern.

Abgesehen vom FDZ DIPF haben alle konsultierten Infrastruktureinrichtungen eine Langzeitarchivierungsstrategie implementiert bzw. ein entsprechendes technisches Konzept entwickelt, wobei sich sowohl Zeiträume als auch Archivierungsstrategien unterscheiden: Am FDZ Ruhr am RWI spricht man von mindestens zehn Jahren, datorium archiviert Forschungsdaten für mindestens 25 Jahre und IANUS beabsichtigt, Daten so lange zu archivieren, wie es die vorhandenen Ressourcen zulassen. Die FDZs Ruhr am RWI und IANUS priorisieren dabei eine migrationsbasierte Langzeitarchivierungsstrategie, um der Gefahr der Dateiosoleszenz vorzubeugen und somit die technische Nachnutzbarkeit der Datensätze für möglichst lange Zeiträume zu gewährleisten. Das Datenrepositorium datorium verändert die Daten nicht, sondern betreibt ausschließlich Bitstream Preservation.³⁴ Eine dauerhafte Nutzbarkeit der Daten kann damit nicht garantiert werden.

4. Vorläufige Empfehlungen

Mit Blick auf die Datenarchive sind trotz der unterschiedlichen disziplinären Anforderungen übereinstimmende Schnittmengen bei den Lösungsansätzen erkennbar. Allgemeine Problemfelder wie die technische Implementierung der Infrastruktur oder der Datenschutz sind für alle relevant, die konkrete Ausführung ist jedoch abhängig von den jeweiligen Forschungsdatentypen und disziplinspezifischen Anforderungen. Auch für die ethnologischen Fächer scheint eine generische Lösung unzureichend zu sein. Weil Ethnologinnen und Ethnologen mit wenig standardisierten Forschungsmethoden in offenen Forschungsprozessen arbeiten und viele Forschungsprobleme erst im Feld gelöst werden können, ist die Heterogenität der Forschungsdaten und Datenformate erheblich.

Benötigt werden deshalb Infrastrukturen, die zum einen technisch flexibel genug sind, die teils sehr komplexen Beziehungen der Daten untereinander adäquat zu dokumentieren und zu kontextualisieren. Zum anderen sollten die Datenarchive unserer Ansicht nach einen hohen Grad an disziplinspezifischer Professionalisierung aufweisen. Denn weil regelrechte Publikationen der vielfach sensiblen Daten aus ethnografischen Forschungen bestenfalls in wenigen Fällen möglich sein werden,

33 Stefan Winkler-Nees, „Stand der Diskussion und Aktivitäten: National,“ in *Langzeitarchivierung von Forschungsdaten: Eine Bestandsaufnahme*, hrsg. Heike Neuroth et al. (Boizenburg: Verlag Werner Hülsbusch, 2012), 35, <http://nbn-resolving.de/urn:nbn:de:0008-2012031401>.

34 Digitale Daten sind Bits, die z.B. auf Festplatten gespeichert werden. Jede Datei besteht aus einer bestimmten Abfolge von Bits, die auch Bitstrom (bitstream) genannt wird. Bitstream Preservation ist eine notwendige Grundlage digitaler Langzeitarchivierung, garantiert aber lediglich den Erhalt der Dateien in der Ausgangsform. Aufgrund von Technologiewandel kann es aber notwendig werden, die Ausgangsdateien z.B. in neuere Formate zu migrieren.

sind die Anforderungen an die Vertrauenswürdigkeit von Datenarchiven entsprechend hoch und vielfältig: Benötigt wird etwa dauerhaft ein umfassendes Rechtemanagement, das nicht intendierte Verwertungen von Daten verhindert, gestufte Zugangsrechte bis hin zur On-Site-Nutzung einschließt und vor allem Zugangskontrolle zuverlässig gewährleistet. Viele Forschende möchten gefragt werden, ob, an wen und wofür ihre Daten zur Nachnutzung freigegeben werden können. Auch aus der Perspektive der Nachnutzung ist die Möglichkeit, Kontakt zu den Primärforschenden aufnehmen zu können, vielen wichtig, vor allem, um zusätzliche Informationen über das Material einzuholen. Ebenso sollten Exit-Strategien bzw. Möglichkeiten diskutiert werden, archivierte Daten wieder zurückzuziehen, etwa wenn sich die Bedingungen ändern, unter denen die Daten übergeben wurden.

Mit Blick auf Daten aus qualitativer Sozialforschung ist aus unserer Sicht Qualiservice derzeit deutschlandweit das einzige Datenarchiv, das Lösungen für diese komplexen Anforderungen erprobt, im Rahmen seiner Möglichkeiten auch praktiziert und geeignet wäre, ethnografische Daten aufzunehmen. Sofern Qualiservice in den Regelbetrieb überführt werden kann, halten wir mit Blick auf die spezifischen Anforderungen ethnografischer Forschung eine weitergehende Kooperation für äußerst wünschenswert. Einer vollständigen Zentralisierung soll hier gleichwohl nicht das Wort geredet werden. Schließlich sind Universitäten und andere Forschungsinstitutionen die Orte, an denen ethnografische Daten entstehen. Sie sind deshalb in der Verantwortung, auch für „schwierige“ Daten, die qualitative Sozialdaten aus Sicht von Datenmanagementprozessen zweifellos sind, Möglichkeiten der langfristigen Archivierung zu schaffen.

Literaturverzeichnis

- Breidenstein, Georg, Stefan Hirschauer, Herbert Kalthoff und Boris Nieswand. *Ethnografie: Die Praxis der Feldforschung*. 2. Auflage. Konstanz: UVK Verlagsgesellschaft mbH, 2015.
- Cliggett, Lisa. „Preservation, Sharing and Technological Challenges of Longitudinal Research in the Digital Age.“ In *eFieldnotes: The Makings of Anthropology in the Digital World*, herausgegeben von Roger Sanjek und Susan W. Tratner, 231–250. Philadelphia: University of Pennsylvania Press, 2016.
- DFG. *Sicherung guter wissenschaftlicher Praxis: Denkschrift*. Ergänzte Auflage. Weinheim: WILEY-VCH Verlag GmbH & Co. KGaA, 2013. Zuletzt geprüft am 28.11.2017. www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf.
- DFG. „Umgang mit Forschungsdaten: DFG-Leitlinien zum Umgang mit Forschungsdaten“. 2015. Zuletzt geprüft am 27.11.2017. http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/index.html.
- Dilger, Hansjörg. „Ethik und Ethnologie: Ethikkommissionen, ethnographisches Arbeiten und Epistemologie – nicht nur in der Medizinethnologie.“ *Medizinethnologie: Körper,*

Gesundheit und Heilung in einer globalisierten Welt (Blog). 16. September 2015. Zuletzt geprüft am 27.11.2017. <http://www.medizinethnologie.net/ethik-und-ethnologie/>.

- Droß, Patrick J. *Kurzstudie: Anforderungen an die Archivierung sozial- und wirtschaftswissenschaftlicher Forschungsdaten*. Berlin: WZB, 2015. Zuletzt geprüft am 28.11.2017. https://sowidatanet.de/images/pdfs/Meldungen/Kurzstudie_Qualitative_Interviews.pdf.
- Emerson, Robert M., Rachel I. Fretz und Linda L. Shaw. *Writing Ethnographic Fieldnotes*. 2. Auflage. Chicago, London: University of Chicago Press, 2011.
- European Commission. „Horizont 2020: Aktualisiertes Arbeitsprogramm unterstützt Wettbewerbsfähigkeit durch offene Wissenschaft“. 2016. Zuletzt geprüft am 28.11.2017. http://europa.eu/rapid/press-release_MEMO-16-2604_de.htm.
- Fabian, Johannes. *Time and the Other. How Anthropology Makes Its Object*. New York: Columbia University Press, 1983.
- Gebel, Tobias, Matthis Grenzer, Julia Kreusch, Stefan Liebig, Heidi Schuster, Ralf Tschewinka, Oliver Watteler und Andreas Witzel. „Verboten ist, was nicht ausdrücklich erlaubt ist: Datenschutz in qualitativen Interviews.“ *Forum: Qualitative Sozialforschung / Forum: Qualitative Social Research* 16, Nr. 2 (2015), Art. 27. <http://nbn-resolving.de/urn:nbn:de:0114-fqs1502279>.
- Huschka, Denis, Hubert Knoblauch, Claudia Oellers und Heike Solga, Hrsg., *Forschungsinfrastrukturen für die qualitative Sozialforschung*. Berlin: SCIVERO Verlag, 2013. Zuletzt geprüft am 13.02.2018. https://www.ratswd.de/dl/downloads/forschungsinfrastrukturen_qualitative_sozialforschung.pdf.
- Imeri, Sabine und Ida Danciu [Mitarb.]. *Open Data. Forschungsdatenmanagement in den ethnologischen Fächern. Auswertung einer Umfrage des Fachinformationsdienstes Sozial- und Kulturanthropologie an der Universitätsbibliothek der Humboldt-Universität zu Berlin 2016, Teil I: Statistiken*. Version 1.0 / Juni 2017. Zuletzt geprüft am 18.06.2018. http://www.evifa.de/cms/fileadmin/uploads/Umfrage_Bericht_Statistiken_1.0_14-06-2017.pdf.
- Kaden, Ben. „Dienste für die dissertationsbegleitende Publikation von Forschungsdaten. Eine Vortragsnachlese.“ *LIBREAS. Library Ideas* (Blog). 07. Juni 2017. Zuletzt geprüft am 27.11.2017. <https://libreas.wordpress.com/2017/06/07/forschungsdaten-2/#more-4917>.
- Knecht, Michi. „Nach Writing Culture, mit Actor-Network: Ethnographie/Praxeographie im Feld der Wissenschafts-, Medizin- und Technikanthropologie.“ In *Europäisch-ethnologisches Forschen: Neue Methoden und Konzepte*, herausgegeben von Sabine Hess, Johannes Moser und Maria Schwertl, 79–106. Berlin: Reimer, 2013.

- Kretzer, Susanne. *Arbeitspapier zur Konzeptentwicklung der Anonymisierung / Pseudonymisierung in Qualiservice*. Bremen: Universität Bremen, 2013. Zuletzt geprüft am 28.11.2017. http://www.qualiservice.org/fileadmin/templates/qualiservice/Anonymisierungskonzept_Arbeitspapier.pdf.
- Lederman, Rena. „Archiving Fieldnotes? Placing ‘Anthropological Records’ Among Plural Digital Worlds.“ In *eFieldnotes: The Makings of Anthropology in the Digital World*, herausgegeben von Roger Sanjek und Susan W. Tratner, 251–271. Philadelphia: University of Pennsylvania Press, 2016.
- Meyermann, Alexia und Maike Porzelt. „Hinweise zur Anonymisierung von qualitativen Daten.“ *forschungsdaten bildung informiert*. Nr. 1 (2014). Zuletzt geprüft am 19.06.2018. <https://www.forschungsdaten-bildung.de/anonymisierung>.
- Peukert, Alexander und Marcus Sonnenberg. „Das Urheberrecht und der Wandel des wissenschaftlichen Kommunikationssystems.“ In *Wissenschaftliches Publizieren. Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung*, herausgegeben von Peter Weingart und Niels Taubert, 211–242. Berlin, Boston: De Gruyter Akademie Forschung, 2016.
- Raitchel, Jürgen. *Quantitative Forschung: Ein Praxiskurs*. 2., durchgesehene Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften, 2008. <https://doi.org/10.1007/978-3-531-91148-9>.
- Stodulka, Thomas. „Feldforschung als Begegnung – Zur pragmatischen Dimension ethnographischer Daten.“ *Sociologus* 64, Nr. 2 (2014): 179–206. <https://doi.org/10.3790/soc.64.2.179>.
- von Unger, Hella, Hansjörg Dilger und Michael Schönhuth. „Ethikbegutachtung in der sozial- und kulturwissenschaftlichen Forschung? Ein Debattenbeitrag aus soziologischer und ethnologischer Sicht.“ *Forum: Qualitative Sozialforschung / Forum: Qualitative Social Research* 17, Nr. 3 (2016), Art. 20. <http://nbn-resolving.de/urn:nbn:de:0114-fqs1603203>.
- Widlok, Thomas. „The Archive Strikes Back: Effects of Online Digital Language Archiving on Research Relations and Property Rights.“ In *Oral Literature in the Digital Age: Archiving Orality and Connecting with Communities*, herausgegeben von Mark Turin, Claire Wheeler, und Eleanor Wilkinson, 3–19. Cambridge: Open Book Publishers, 2013. Zuletzt geprüft am 28.11.2017. <https://www.openbookpublishers.com/reader.php/186?186>.
- Winkler-Nees, Stefan. „Stand der Diskussion und Aktivitäten: National.“ In *Langzeitarchivierung von Forschungsdaten: Eine Bestandsaufnahme*, herausgegeben von Heike Neuroth, Stefan Strathmann, Achim Oßwald, Regine Scheffel, Jens Klump und Jens Ludwig. Boizenburg: Verlag Werner Hülsbusch, 2012. <http://nbn-resolving.de/urn:nbn:de:0008-2012031401>.

- Wissenschaftsrat. *Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020*. Drs. 2359-12. Berlin: WR, 2012. Zuletzt geprüft am 28.11.2017. <https://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>.

Der Chimäre auf der Spur: Forschungsdaten in den Geisteswissenschaften

Fabian Cremer, Max Weber Stiftung

Lisa Klaffki, DARIAH-DE / Herzog August Bibliothek Wolfenbüttel

Timo Steyer, Forschungsverbund Marbach Weimar Wolfenbüttel / Herzog August Bibliothek Wolfenbüttel

Zusammenfassung:

Forschungsdaten nehmen in den Geistes- und Kulturwissenschaften eine immer wichtigere Rolle im Forschungsprozess ein. Doch dieser Vorgang ist nicht frei von Widersprüchen und Missverständnissen – das aus den Naturwissenschaften stammende Konzept „Forschungsdaten“ kann nur bedingt auf traditionelle Arbeitsverfahren in den Geisteswissenschaften übertragen werden. Der vorliegende Beitrag thematisiert anhand der Spezifika geisteswissenschaftlicher Forschungsdaten das Spannungsverhältnis zwischen Relevanz und Wahrnehmung bei den Forschenden sowie zwischen Konzeption und Umsetzung im Forschungsalltag.

Summary:

Research data is getting more and more important in the humanities and cultural studies and now plays an important role in the research process. But contradictions and misunderstandings are also part of this development. The concept of research data comes from the sciences, and can only be transferred to the traditional scholarly methods of the humanities to a limited degree. Based on the specifics of research data in the humanities, the article deals with the tensions between relevance and familiarity among the researchers as well as between the conception and implementation in the research process.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S142-162>

Autorenidentifikation: Cremer, Fabian: GND: 14061773, ORCID: <https://orcid.org/0000-0001-8251-9727>; Klaffki, Lisa: GND: 1041562500, ORCID: <https://orcid.org/0000-0002-8791-7422>;

Steyer, Timo: GND: 1053806175, ORCID: <http://orcid.org/0000-0003-0218-2269>

Schlagwörter: Forschungsdaten, Digital Humanities

1. Genese mit Widersprüchen

In einem Blogbeitrag stellte Thomas Stäcker die Frage, ob Forschungsdaten in den Geisteswissenschaften “[...] lediglich eine Chimäre, ein von den Technik- und Naturwissenschaften geborgter Begriff [sein]?”¹

1 Thomas Stäcker, „Noch einmal: Was sind geisteswissenschaftliche Forschungsdaten?“, *DHd-Blog*, 6. Dezember 2015, zuletzt geprüft am 16.01.2018, <http://dhd-blog.org/?p=5995>. Der Blogbeitrag war eine Replik auf ein von Peter Andorfer verfasstes DARIAH-DE Working Paper. Vgl. Peter Andorfer, *Forschungsdaten in den (digitalen) Geisteswissenschaften. Versuch einer Konkretisierung*, DARIAH-DE Working Papers 14 (Göttingen: DARIAH-DE, 2015), <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-7-2>.

Gut zwei Jahre nach der Formulierung dieser Frage lässt sich eine zunehmende Relevanz von Forschungsdaten in geisteswissenschaftlichen Forschungsprozessen feststellen, die sich auch in einer Institutionalisierung in Form von Arbeitsstellen und Datenzentren ausdrückt.² Dabei spielen die Aspekte der Forschungsdatenpublikation und der Anreicherung von Forschungsdaten hinsichtlich der weiteren Nachnutzung eine zentrale Rolle. Das Chimärenhafte bleibt diesem Begriff jedoch weiterhin verhaftet, nicht länger als Trugbild, dafür als vielgestaltiges Mischwesen erscheinen die unterschiedlichen Praktiken und Perspektiven zu Forschungsdaten in den Geisteswissenschaften.³ In dieser noch nicht abgeschlossenen Begriffsgenese wird der Begriff Forschungsdaten auf unterschiedliche Qualitäts- und Aggregatsstufen von Daten, angefangen bei Grunddaten wie Digitalisaten oder maschinell erstellten Volltexten bis zu nach komplexen Richtlinien erschlossenen Datensammlungen oder Datenrepräsentation bzw. Datenbanken, angewendet und ist nicht frei von Widersprüchen und Akzeptanzproblemen.⁴

Während einige Wissenschaftler/innen vor allem im Bereich der digitalen Geisteswissenschaften (Digital Humanities) von einem Paradigmenwechsel hin zu maschinenlesbaren, strukturierten und frei nachnutzbaren Daten sprechen, hat der/die eher traditionell hermeneutisch arbeitende Geisteswissenschaftler/in oft Vorbehalte gegenüber der Publikation von Daten, die als Arbeitsmittel auf dem Weg zur Monographie betrachtet werden, wie z.B. Exceltabellen oder Filemakerdatenbanken sowie dem zeitintensiven Anreichern dieser Arbeitsdaten. Besonders in diesem Kontext sind im Gegensatz zu den in standardisierten Digitalisierungsworkflows generierten Forschungsdaten die skizzierten Widersprüche virulenter und die Positionen noch nicht konsensfähig.⁵ Auch kann generell hinterfragt

- 2 So widmet sich das Historische Datenzentrum Sachsen-Anhalt an der Universität Halle-Wittenberg dem geisteswissenschaftlichen Forschungsdatenmanagement. Vgl. <http://www.geschichte.uni-halle.de/struktur/hist-data/>, zuletzt geprüft am 16.01.2018. Auch wäre in diesem Kontext das Projekt „Geisteswissenschaftliche Infrastruktur für Nachhaltigkeit“ (gwin) zu nennen, welches im Rahmen des Zentrums für nachhaltiges Forschungsdatenmanagement eine Infrastruktur für Forschungsdaten und Forschungsanwendungen an der geisteswissenschaftlichen Fakultät der Universität Hamburg erstellt. Vgl. <https://www.fdm.uni-hamburg.de/ueber-uns/gwin.html>, zuletzt geprüft am 16.01.2018. Eine Bündelung der Initiativen von geisteswissenschaftlichen Datenzentren findet in der Arbeitsgruppe Datenzentren des DHD-Vereins statt, zuletzt geprüft am 16.01.2018, <https://dig-hum.de/ag-datenzentren>.
- 3 Vgl. Patrick Sahle und Simone Kronenwett, „Jenseits der Daten: Überlegungen zu Datenzentren für die Geisteswissenschaften am Beispiel des Kölner 'Data Center for the Humanities',“ *LIBREAS. Library Ideas*, 23 (2013), <http://nbn-resolving.de/urn:nbn:de:kobv:11-100212726>.
- 4 Michael Knoche konstatiert in diesem Kontext einen Wildwuchs an unterschiedlichen Initiativen und das Fehlen klarer Vorstellungen wie Wissenschaftler/innen beim Thema Forschungsdaten von bibliothekarischer Seite unterstützt werden sollen. Vgl. Michael Knoche, *Die Idee der Bibliothek und ihre Zukunft* (Göttingen: Wallstein, 2017), 22.
- 5 Die höhere Standardisierung bei institutionell erhobenen Forschungsdaten, wie z.B. Digitalisaten, wird durch die standardbildenden Förderrichtlinien gestärkt. Seit 2009 empfiehlt die DFG einen fachspezifischeren Umgang mit Forschungsprimärdaten und deren Organisation, eine standardisierte Speicherung mit Rechteverwaltung, den freien Zugang, die Auszeichnung mit Metadaten, sowie Verfahren zur Qualitätssicherung, vgl. Deutsche Forschungsgemeinschaft, Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme, Unterausschuss für Informationsmanagement, „Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten,“ 2009, zuletzt geprüft am 19.02.2018, http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901.pdf; DFG, „Antragstellung,“ zuletzt geprüft am 19.02.2018, http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf; DFG, „Leitlinien zu den Forschungsdaten,“ 2015, zuletzt geprüft am 19.02.2018, http://www.dfg.de/foerderung/info_wissenschaft/2015/info_wissenschaft_15_66/; Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen, „Grundsätze zum Umgang mit Forschungsdaten,“ 2010, <http://dx.doi.org/10.2312/ALLIANZOA.019>.

werden, inwieweit ein Bewusstsein für die Wichtigkeit von Daten und die Aufwendung eigener Zeit- und Arbeitsressourcen für deren Verarbeitung in den Geisteswissenschaften etabliert ist.⁶

Hinzu kommt die unklare infrastrukturelle Verantwortlichkeit für die Erwerbung, Präsentation, Pflege und Archivierung von Forschungsdaten. Das Label Forschungsdaten wird dabei aufgrund seiner Aktualität auf bereits etablierte Konzepte, wie z.B. auf Meta- oder Erschließungsdaten, übertragen oder inflationsartig für jede Art von Daten verwendet. Cum grano salis: Während für die einen Forschungsdaten per se auch Forschungsergebnisse sind, sind Forschungsdaten für andere lediglich relevante, aber nicht publikationsfähige Nebenprodukte ihres Forschungsvorhabens – ihre Aufwertung zu Forschungsdaten erscheint als Trugbild des digitalen Wandels. Entscheidend ist dabei immer die Frage nach dem Erkenntnisgewinn, der abgesehen von einem häufig theoretischen Diskurs über Nachnutzungsszenarien noch nicht beantwortet ist.⁷ Der vorliegende Beitrag versteht sich dabei nicht als Versuch, eine allgemein gültige Definition für den Begriff der Forschungsdaten zu liefern, sondern möchte vielmehr die Spezifika geisteswissenschaftlicher Forschungsdaten herausarbeiten und darauf aufbauend der Frage nachgehen, warum trotz der Relevanz von Forschungsdaten ein Spannungsverhältnis zwischen Theorie (Anforderungen) und Praxis (Verwendung) existiert.⁸

2. Kreation statt Generierung

Mit der zunehmenden Etablierung der Digital Humanities in den Geistes- und Kulturwissenschaften bilden bewusst als solche wahrgenommene und gezielt erzeugte Forschungsdaten eine neue Facette innerhalb der Geistes- und Kulturwissenschaften. Die auf eine Forschungsfrage zielorientierte Datengenerierung wird zunehmend als notwendiger Teil für die Erlangung von neuen Erkenntnissen definiert und rückt somit auch immer stärker in das Bewusstsein der Forschungscommunity.⁹ Zwar bildet die massenhafte Digitalisierung von Objekten des in den Gedächtnisinstitutionen aufbewahrten kulturellen Erbes weiterhin einen signifikanten Anteil der generierten Daten in den Geistes- und Kulturwissenschaften, aber digital innerhalb eines Forschungsprozesses erzeugte Daten (digital-born) ergänzen zunehmend das Portfolio der Datenerzeugung.¹⁰ Während digital-born data also forschungs-

6 Vgl. Christoph Schöch, „Big? Smart? Clean? Messy? Data in the Humanities,“ *Journal of Digital Humanities* 2, Nr. 3 (2013), zuletzt geprüft am 24.02.2018, <http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/>.

7 Vgl. Constanze Baum und Thomas Stäcker, „Die Digital Humanities im deutschsprachigen Raum. Methoden – Theorien – Projekte,“ in *Grenzen und Möglichkeiten der Digital Humanities*, hrsg. Constanze Baum und Thomas Stäcker, Sonderband der Zeitschrift für digitale Geisteswissenschaften 1 (Wolfenbüttel: Herzog August Bibliothek, 2015), http://dx.doi.org/10.17175/sb001_023.

8 Hubertus Kohle bezeichnet die derzeitige Publikationstätigkeit von Forschungsdaten in den Geisteswissenschaften als „sporadisch“. Vgl. Hubertus Kohle, „Digitales Publizieren,“ in *Digital Humanities: Eine Einführung*, hrsg. Fotis Jannidis, Hubertus Kohle und Malte Rehbein (Stuttgart: J. B. Metzler, 2017), 201.

9 In diesem Kontext kann auch der Titel „Von Daten zu Erkenntnissen: Digitale Geisteswissenschaften als Mittler zwischen Information und Interpretation“ der Jahrestagung 2015 des Vereins Digital Humanities im deutschsprachigen Raum gewertet werden. Die Tagung fand vom 23.02. bis zum 27.02.2015 an der Universität Graz statt. Vgl. Maik Fiedler und Andreas Weiß, „Von Daten zu Erkenntnissen: Digitale Geisteswissenschaften als Mittler zwischen Information und Interpretation. DHd-Jahrestagung 2015,“ *H-Soz-Kult* (Blog), 6. Juli 2015, zuletzt geprüft am 13.02.2018, <http://www.hsozkult.de/conferencereport/id/tagungsberichte-6059>.

10 Gemessen am Gesamtbestand ist die Anzahl an digitalisierten Quellen noch immer gering. Leider fehlt es an einer koordinierten nationalen Strategie für die Digitalisierung des gesamten kulturellen Erbes. Trotz übergreifender Projekte wie z. B. dem VD 17 stellen viele Digitalisierungsmaßnahmen Einzelvorhaben dar. Vgl. Lisa Klaffki, Stefan Schmunk und Thomas Stäcker, *Stand der Kulturgutdigitalisierung in Deutschland: Eine Analyse und Handlungsvorschläge*

getrieben (research-driven) erzeugt werden, sind vor allem die Surrogate von historischen Beständen von Bibliotheken und Archiven aus Gründen der nachhaltigen Erschließung produziert (curation-driven). Das aus den Naturwissenschaften bekannte Konzept der Kategorisierung der Daten nach Primär- und Sekundärdaten greift in den Geisteswissenschaften nicht, da auch innerhalb desselben Forschungsprozesses Sekundärdaten wieder zu Primärdaten werden; dies trifft vor allem auch für die Nachnutzung von Forschungsdaten zu.¹¹ Ebenfalls einen Unterschied zu den naturwissenschaftlichen Fächern bildet das Verhältnis von automatisierten Verfahren zu manuellen Erfassungs- und Erschließungsarbeiten. Trotz erheblicher Fortschritte in den Bereichen der Volltexterkennung, der automatisierten Erschließung und Annotation der Daten sind bei der Erhebung von hochqualitativen wissenschaftlichen Forschungsdaten in den Geisteswissenschaften nach wie vor größere manuelle Arbeitsanteile vonnöten. Zum einen ist dies durch die hochspezialisierten Fragestellungen bedingt, zum anderen durch die Unikalität und die Spezifika der historischen Überlieferung.¹² Die Mehrzahl der geisteswissenschaftlichen Projekte basiert nicht auf großen Datenmengen, sondern auf sehr kleinen und komplexen Datensammlungen, die die Historizität der zugrundeliegenden Quellen wiedergeben.¹³

So wurde im Projekt Frühneuzeitliche Gelehrtenbibliotheken des Forschungsverbunds Marbach Weimar Wolfenbüttel (MWW) die Erschließung der Bibliothek des Mathematikers und Chiliasten Benedikt Bahnsen über die Auswertung des Auktionskatalogs der Privatbibliothek durch eine weitestgehend manuelle Transkription der historischen Vorlage anhand eines definierten Datenmodells vorgenommen.¹⁴

des DARIAH-DE Stakeholdergremiums ‚Wissenschaftliche Sammlungen‘, DARIAH-DE Working Papers Nr. 26 (Göttingen: DARIAH-DE, 2018), <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2018-1-3>; sowie das zugehörige Memorandum: Lisa Klaffki, „Memorandum: Digitalisierung von Kulturgut,“ *DHd-Blog*, 17. Mai 2018, zuletzt geprüft am 05.06.2018, <https://dhd-blog.org/?p=9849>.

- 11 Es wird daher empfohlen, in den Geisteswissenschaften keine Unterscheidung zwischen Primär- und Sekundärdaten vorzunehmen und allgemein von Forschungsdaten zu sprechen.
Vgl. Jens Klump, „Digitale Forschungsdaten,“ in *Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*, hrsg. Heike Neuroth et al., Version 2.3 (Göttingen: nestor, 2010), Kap. 17.10, 17:105, <http://nbn-resolving.de/urn:nbn:de:0008-2010071949>; vgl. Johanna Puhl et al., *Diskussion und Definition eines Research Data LifeCycle für die digitalen Geisteswissenschaften*, DARIAH-DE Working Papers Nr. 11 (Göttingen: DARIAH-DE, 2015), 9, <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-4-4>.
- 12 Exemplarisch können die Probleme bei der Volltexterkennung von frühneuzeitlichen Leichenpredigten angeführt werden. Vgl. Maria Federbusch, „OCR für Drucke der Frühen Neuzeit? Erfahrungen und Perspektiven am Beispiel von Funeralschriften,“ *Bibliotheksdienst* 49, Nr. 7 (2015): 713–724; Thomas Stäcker, „Erfahrungsbericht Helmstedter Drucke Online an der Herzog August Bibliothek Wolfenbüttel,“ in *Volltext via OCR: Möglichkeiten und Grenzen. Test-szenarien zu den Funeralschriften der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz*, hrsg. Maria Federbusch und Christian Polzin, Beiträge aus der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz 43 (Berlin: Staatsbibliothek zu Berlin, 2013), 123–135.
- 13 Generell kann in den Geisteswissenschaften trotz der Massendigitalisierung nicht von Big Data gesprochen werden. Vgl. Schöch, „Data in the Humanities“.
- 14 Vgl. Hartmut Beyer et al., „Bibliotheken im Buch: Die Erschließung von privaten Büchersammlungen der Frühneuzeit über Auktionskataloge,“ in *Kodikologie und Paläographie im digitalen Zeitalter 4* (Codicology and Palaeography in the Digital Age), hrsg. Hannah Busch, Franz Fischer und Patrick Sahle (Norderstedt: Books on Demand, 2017), 43–70, <http://nbn-resolving.de/urn:nbn:de:hbz:38-77794>.

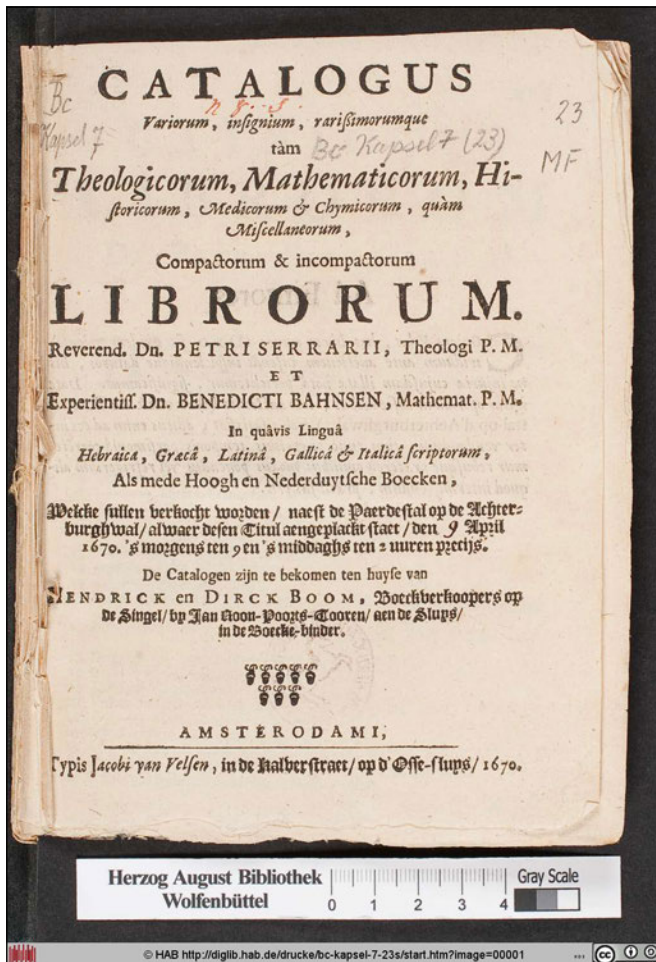


Abb. 1: Titelblatt des Auktionskatalogs der Bibliothek von Benedikt Bahnsen.

Ebenso erfolgte die Anreicherung der erhobenen biobibliographischen Daten mit Normdaten in gemeinsamer Arbeit von Wissenschaftler/in und Bibliothekar/in. Erst im Zuge der Auswertung und Analyse der Daten wurden vermehrt automatisierte Verfahren angewandt. Ferner wurden die Druckorte der verzeichneten Bücher mittels des DARIAH-DE Geobrowser in Relation zum Druckjahr visualisiert sowie ausgewählte Inhalte des Kataloges, wie z.B. Autoren, inhaltliche Kategorien oder Hauptsprachen, mittels unterschiedlicher Visualisierungsvarianten von Wortwolken bis Kreisdiagrammen abgebildet.¹⁵

15 Vgl. die Rekonstruktion der Bibliothek von Benedict Bahnsen, zuletzt geprüft am 19.02.2018, <http://dev.hab.de/auktionskataloge/bahnsen/bahnsen-catSubjectFormat.html>.

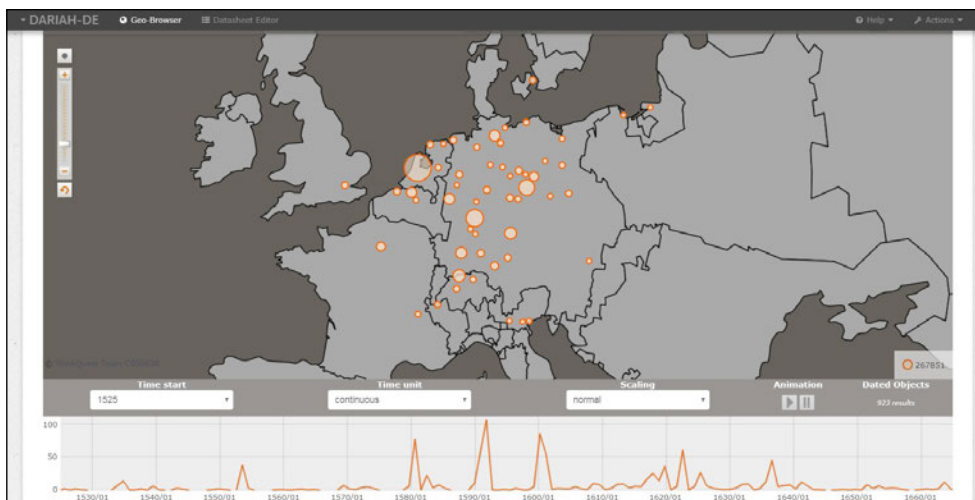


Abb. 2: Visualisierung der Druckorte / Druckjahre des Auktionskatalogs von Benedikt Bahsen im DARIAH-DE Geobrowser.

Das Beispiel zeigt, dass die Erstellung eines Forschungsdatensets eher eine domänenspezifischen Kreation denn ein automatisierter Datengenerierungsprozess ist. Gerade durch diesen Zuschnitt wird die Bedeutung von wissenschaftlichen Bibliotheken als Bereitsteller der Quellen, Kompetenzzentrum für deren Erschließung und als Anbieter für technische Lösungen – in summa die Bibliothek als traditionelle Infrastruktur für die Geisteswissenschaften – auch bzw. insbesondere im digitalen Forschen gestärkt.¹⁶

3. Unbehagen am Trend

Die Vielfalt und Heterogenität von geisteswissenschaftlichen Forschungsdaten korrespondiert zwangsläufig mit einer sehr weiten Definition, wie sie z.B. im DARIAH-DE Working Paper zum Research Data LifeCycle zu finden ist: „Unter digitalen geistes- und kulturwissenschaftlichen Forschungsdaten werden innerhalb von DARIAH-DE all jene Quellen/Materialien und Ergebnisse verstanden, die im Kontext einer geistes- und kulturwissenschaftlichen Forschungsfrage gesammelt, erzeugt, beschrieben und/oder ausgewertet werden und in maschinenlesbarer Form zum Zwecke der Archivierung, Zitierbarkeit und zur weiteren Verarbeitung aufbewahrt werden können.“¹⁷ Während die Funktionen des Sammelns, Beschreibens und Auswertens etablierte Vorgänge in den Geistes- und Kulturwissenschaften darstellen und keineswegs ein Spezifikum oder ein Resultat des digital turn sind, stellen die Aspekte der Zitierbarkeit, der Archivierung und des Teilens unter dem Paradigma der Maschinenlesbarkeit das Innovationspotenzial der Forschungsdaten heraus.¹⁸ Während die für den

16 Vgl. Petra Maier, *Digital Humanities und Bibliothek als Kooperationspartner*, DARIAH-DE Working Papers Nr. 19 (Göttingen: DARIAH-DE, 2016), 9, <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2016-5-6>.

17 Johanna Puhl et al., *Diskussion*, 14.

18 Vgl. Angelina Kraft, „Die FAIR Data Prinzipien für Forschungsdaten,“ TIB-Blog – Weblog der Technischen Informationsbibliothek (TIB), 12. September 2017, zuletzt geprüft am 24.01.2018, <https://blogs.tib.eu/wp/tib/2017/09/12/>

Erkenntnisgewinn notwendigen Zwischenprodukte, wie z.B. tabellarische Auswertungen, Arbeitsdatenbanken, Literaturlisten oder Exzerpte, nicht oder nur rudimentären Eingang in das schriftliche Ergebnis des Forschungsprozesses fanden, ist es nun technisch möglich, auch diese Zwischenprodukte zu publizieren und sie z.B. in Form einer enhanced publication mit dem Resultat zu verbinden.¹⁹ Die Publikation von Forschungsdaten würde den Forschungsweg transparenter und nachvollziehbarer machen, gleichzeitig das Anknüpfen an das Forschungsvorhaben erleichtern.

Trotz dieser Vorteile ist nicht abzusehen, ob sich dieser Trend abseits der Digital Humanities in den Geistes- und Kulturwissenschaften durchsetzt; scheint er für viele Geisteswissenschaftler/innen doch ihrem Verständnis von geisteswissenschaftlicher Forschung zu widersprechen: Angefangen bei dem ungewohnten Vorgang der zügigen Publikation von Zwischenergebnissen, der Transparenz des eigenen Forschungsprozesses sowie dem Unbehagen von technischen Abhängigkeiten bis zur Beurteilung der Datenqualität, der Reputation für die Publikation von Forschungsdaten bis hin zu der Bearbeitung durch Dritte im Zuge der Anreicherung und des Nachweises von Forschungsdaten. Dabei werden häufig die unterschiedlichen Qualitätsstufen von Datensets nicht ausreichend thematisiert bzw. gekennzeichnet, denn auch unstrukturierte und fehlerbehaftete Daten können für weitere Forschungsfragen von Relevanz sein bzw. als Grundlage für elaborierte Forschungsvorhaben dienen. Hier bestehen bereits Aktivitäten, wie z. B. die Kennzeichnung der Genauigkeit von OCR-Verfahren oder die Möglichkeit für Nutzer/innen, sich an der Optimierung von publizierten Texten zu beteiligen.²⁰ Gerade durch den Prozess des automatischen Indexierens werden Bibliotheken gefragt sein, sich in diesen Fragen zu positionieren.

4. Kosmos im Kontext

Die Ansprüche an geeignete Publikationsformate für die Repräsentation geisteswissenschaftlicher Daten sind hoch: Sie sollen die Heterogenität und Multimedialität, aber auch die Historizität sowie die Kontextualisierung der Untersuchungsobjekte adäquat darstellen können. In den Anfängen des „elektronischen Publizierens“ in den Geisteswissenschaften wurden vor allem die multimedialen und interaktiven Möglichkeiten der Webtechnologien als Verbesserung und Anreiz wahrgenommen.²¹ Als Resultat entstanden Webseiten, die Forschungsergebnisse und Untersuchungsgegenstände gleichzeitig repräsentieren und kontextualisieren können. Diese digitalen Forschungsdatensammlungen (*Thematic Research Collections*) ersetzen nicht klassische Publikationsformate wie Buch und Artikel,

[die-fair-data-prinzipien-fuer-forschungsdaten/](#).

- 19 Vgl. Andreas Degkwitz, „Enhanced Publications Exploit the Potential of Digital Media,“ in *Evolving Genes of ETDs for Knowledge Discovery. Proceedings of ETD 2015*, 18th International Symposium on Electronic Theses and Dissertations, 51-59, zuletzt geprüft am 19.02.2018, https://www.researchgate.net/profile/Andreas_Degkwitz/publication/283783359_Enhanced_Publications_Exploit_the_Potential_of_Digital_Media/links/5646fe2908ae9f9c13e8c4fa/Enhanced-Publications-Exploit-the-Potential-of-Digital-Media.pdf.
- 20 Vgl. das Tool DTAQ (Deutsches Textarchiv Qualitätssicherung); „Kollaborative Qualitätssicherung im Deutschen Textarchiv,“ zuletzt geprüft am 19.02.2018, <http://www.deutschestextarchiv.de/dtaq/about>.
- 21 Zu den Anfängen des digitalen Publizierens (dort noch elektronisches Publizieren genannt) in den Geisteswissenschaften ist das Positionspapier der DFG lesenswert: DFG, *Elektronisches Publizieren* (Bonn: DFG, März 2005), zuletzt geprüft am 19.02.2018, http://www.dfg.de/download/pdf/foerderung/programme/lis/pos_papier_elektron_publizieren_0504.pdf; hierzu auch die Empfehlungen des Unterausschusses für elektronisches Publizieren und des Ausschusses für wissenschaftliche Bibliotheken und Informationssysteme, Juni 2006, zuletzt geprüft am 19.02.2018, http://www.dfg.de/download/pdf/foerderung/programme/lis/elektr_publizieren.pdf.

sondern bilden eine eigenständige Publikationsform.²² Denn die Webtechnologien erlauben es, die Grundprinzipien geisteswissenschaftlicher Methodik – nach John Unsworth: Annotation, Vergleich, Referenz, Auswahl, Illustration und Repräsentation²³ – in unmittelbarer Form darzustellen, während dies in reinen Texten nur mittelbar durch eine sprachliche Beschreibung möglich ist.

Die Webseite mit zugehöriger Datenbank als wissenschaftliches Publikationsformat und komplexes Informationsobjekt zugleich impliziert, dass nicht nur die hinterlegten und präsentierten Daten, sondern die Präsentationsschicht und Softwarearchitektur selbst Informationsträger wissenschaftlicher Erkenntnis sind. Dies erschwert wesentlich die für eine langfristige Erhaltung notwendige Trennung zwischen der webbasierten Präsentations- und einer dateibasierten Archivierungsschicht.²⁴ Die komplexen Webseiten haben sich damit zu echten Sorgenkindern für die Informationsinfrastruktureinrichtungen entwickelt.²⁵ Diesen Herausforderungen lässt sich durch Investitionen in Forschungsdatenmanagement bei Forschenden, Förderern und Informationsinfrastruktureinrichtungen begegnen und hier sind auf allen Ebenen deutliche Anzeichen der Bewusstseinsbildung und Handlungswille zu konstatieren.²⁶

Die zahlreichen Anforderungen des modernen Forschungsdatenmanagement wurden bei der Online-Edition der "Teutschen Academie" von Joachim von Sandrart (sandrart.net) nahezu vorbildlich erfüllt: standardisierte Datenmodelle (XML/TEI, CIDOC-CRM), persistente Identifikatoren, Normdaten (GND, TGN, etc.), offene Schnittstellen und Webservices (REST-API, LOD/RDF, Beacon), offene Lizenzen (ODbL, CC-BY-NC-SA).²⁷ Die Textdaten liegen zudem in einem Repositorium.²⁸ Dennoch vermag die Datenschicht nicht, die komplexen Funktionalitäten des nutzerfreundlich und effizient gestalteten Webinterface zu erhalten: Register, facettierte Suche, Links und interaktive Verknüpfungen sowie die einzelnen Seiten zu Personen, Orten, Kunstwerken, bibliographische Nachweise und den Quellen, die als aufbereitete Dossiers das komplexe Beziehungsgeflecht in den Daten erst rezipierbar

22 Vgl. Carole L. Palmer, „Thematic Research Collections,“ in *A Companion to Digital Humanities*, hrsg. Susan Schreibman, Ray Siemens und John Unsworth (Oxford: Blackwell 2004), Kap. 24, zuletzt geprüft am 19.02.2018, <http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-4-5&toc.depth=1&toc.id=ss1-4-5&brand=default>.

23 John Unsworth, *Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?* Vortrag, Humanities Computing: Formal Methods, Experimental Practice, King's College London, 13. Mai 2000, zuletzt geprüft am 19.02.2018, <http://www.iath.virginia.edu/~jmu2m/Kings.5-00/primitives.html>.

24 Wolfgang Pemppe, „Geisteswissenschaften,“ in *Langzeitarchivierung von Forschungsdaten: Eine Bestandsaufnahme*, hrsg. Heike Neuroth et al. (Göttingen: Universitätsverlag Göttingen / Boizenburg: Verlag Werner Hülsbusch, 2012), 141, <http://nbn-resolving.de/urn:nbn:de:0008-2012031401>.

25 Vgl. hierzu Stefan Buddenbohm, Claudia Engelhardt und Ulrike Wuttke, „Angebotsgenese für ein geisteswissenschaftliches Forschungsdatenzentrum,“ *Zeitschrift für digitale Geisteswissenschaften* (2017), http://dx.doi.org/10.17175/2016_003.

26 Zu Maßnahmen der Infrastruktureinrichtungen siehe Anm. 2, aber auch aus der Wissenschaft werden Standards definiert. So hat etwa das Fachkollegium Literaturwissenschaft der DFG Förderkriterien aufgestellt, zuletzt geprüft am 19.02.2018, http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/foerderkriterien_editionen_literaturwissenschaft.pdf.

27 Joachim von Sandrart, *Teutsche Academie der Bau-, Bild- und Mahlerey-Künste, Nürnberg 1675-1680, wissenschaftlich kommentierte Online-Edition*, hrsg. Thomas Kirchner et al., 2008-2012, zuletzt geprüft am 19.02.2018, <http://ta.sandrart.net>.

28 Joachim von Sandrart, *L'Academia Todeasca. della Architectura, Scultura & Pittura: Oder Teutsche Academie der Edlen Bau- Bild- und Mahlerey-Künste. Bd. 1,1. Nürnberg: 1675*. Deutsches Textarchiv, <http://nbn-resolving.de/urn:nbn:de:kobv:b4-30378-3>.



Startseite

Die Deutsche Academie im Volltext
Gliederung der Deutschen Academie
Übersetzungen

Suche
Personen
Orte
Kunstwerke
Bibliographie

Editionsinformationen
Projektinformationen
Herausgeber
Kontakt/Impressum



Permanente URL:
<http://ta.sandrart.net/artwork-4842>

Mischwesen (TA 1680, Tafel CC, 5)

Basis-Daten

Datierung
1680

Material/Technik
Radierung

Signatur und Bezeichnung
ed Sandr:

loh Meyer fecit.

Abbildung(en) in der »Teutschen Academie«
[TA 1680, Iconologia Deorum, Tafel CC. \(nach S. 196\)](#)

Erwähnungen in der Teutschen Academie

»5. Chimaera, oder Amuletum.«
[TA 1680, Iconologia Deorum, S. 196](#)

»Das nachfolgende Stück soll ein Chimaera seyn: das ich aber viel eigentlicher ein Amuletum wider Zauberey und alles Böse/...«
[TA 1680, Iconologia Deorum, S. 196](#)

»... einige unversehene Schrecken Panici genennet worden/ so von diesem Gott sollen hergerühret haben. Diese Abbildung aber ist aus einem alten Niccolo-Stein abgesehen worden.«
[TA 1680, Iconologia Deorum, S. 197](#)

Kommentare

Als graphische Vorlage konnte ein Blatt in [Agostini, Gemme antiche](#) identifiziert werden, welches das Mischwesen gegenseitig abbildet.
► [Carolin Ott, 14.06.2011](#)

Wurde radiert von
– Meyer, Johannes <der Jüngere>

Kunstwerk-Bezüge [Symbolansicht](#)

Ist Teil von
– Meyer: [Tafel mit sechs Medaillons \(TA 1680, Tafel CC\)](#) *Graphik*

Steht in Zusammenhang mit
– Unbekannter antiker Künstler: [Gemme mit Darstellung eines Mischwesens](#) *Kunstgewerbe*

Geschaffen nach
– Galestruzzi: [Mischwesen 1 \(Agostini, Gemme antiche\)](#) *Graphik*

Abb. 3: Entitätsseite eines Kunstwerks in sandrart.net

machen (Abb. 3). Die Webseite funktioniert so als Benutzerschnittstelle, die nicht nur präsentiert, sondern als „virtuelle Forschungsplattform“ die Edition unmittelbar nutzbar macht. Die Webseite mit Oberfläche und Interface wird glücklicherweise von einer wissenschaftlichen Bibliothek betreut und gepflegt.²⁹ Die Projektverantwortlichen haben auf die Herausforderung der langfristigen Nutzbarkeit vorbildliche Antworten gefunden,³⁰ jedoch muss betont werden, dass das ehemalige DFG-Projekt

29 Die Herzog August Bibliothek Wolfenbüttel hat die Digitalisierung der Bände der „Teutschen Academie“ übernommen und sorgt für die Langzeitverfügbarkeit der Edition. „Kooperationen,“ zuletzt geprüft am 19.02.2018, <http://www.sandrart.net/de/kooperationen/>.

30 Patrick Sahle, „Review von ‚Sandrart.net,‘“ *RIDE* 1 (2014): Abs. 11, <https://dx.doi.org/10.18716/ride.a.1.5>.

technologische und informationswissenschaftliche Verfahren einsetzen konnte, die erst heute als zeitgemäße Standards gelten. In der Regel müssen sich Forschungsprojekte aber auf die verfügbaren Best Practices beschränken und können selten Ressourcen in technologische oder informationswissenschaftliche Innovationen investieren.

Sandrat.net demonstriert, dass die Verknüpfung von Texten und Objekten in der Präsentationsebene der Webseiten keineswegs das Resultat fehlender technischer und informationswissenschaftlicher Kompetenz sind, sondern vielmehr Ausdruck geisteswissenschaftlicher Methodik bei der durch Kontextualisierung komplexe Zusammenhänge analysiert werden. Weder lässt sich also eine einfache Trennung zwischen Präsentations- und Datenschicht etablieren; auch der Informations- und Nachnutzungsgehalt einer Datenschicht steht damit in Frage. Auswahl und Aufbereitung der Daten sind in den Geisteswissenschaften an spezifische Fragestellungen, bestimmte disziplinabhängige Methoden und theoretische Perspektiven gebunden. Eine unmittelbare Nachnutzung ist daher nicht ohne Weiteres möglich.³¹ Für nachnutzbare, maschinenlesbare Forschungsdaten bedarf es in den Geisteswissenschaften daher auch geeigneter Repräsentations- und Publikationsformate, die die notwendige Dokumentation und Kontextualisierung der Daten transportieren können. Die webbasierten *Thematic Research Collections* weisen hier den Weg: Sie sind nicht als Nachweis, sondern als nutzbare Ressource konzipiert und gestaltet.

5. Raum oder Wegweiser

Die thematische Vielfalt der Forschungsprojekte bedingt auch stets eine technische Vielfalt der erstellten Ressourcen, aus denen heterogene Datensammlungen entstehen.³² Momentan werden diese, wenn überhaupt, häufig separat von der „eigentlichen Publikation“ in verschiedene Repositorien hochgeladen. Als Repositorien werden derzeit verstärkt domänenspezifische³³ Lösungen umgesetzt, Publikationen zu verwandten Themen und mit einem ähnlichen Publikum befinden sich dann in unmittelbarer Nachbarschaft und können entdeckt werden – ein erster Schritt „vom Datencontainer zum Diskursraum“.³⁴ Für Forschungsdatenrepositorien lassen sich die Infrastrukturen

31 Vgl. Sahle und Kronenwett, *Jenseits der Daten*, 82.

32 Vgl. Tobias Gradl, Andreas Henrich und Christoph Plutte, „Heterogene Daten in den Digital Humanities: Eine Architektur zur forschungsorientierten Föderation von Kollektionen“ in *Grenzen und Möglichkeiten der Digital Humanities*, hrsg. Constanze Baum und Thomas Stäcker. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1 (Wolfenbüttel: Herzog August Bibliothek und Forschungsverbund Marbach Weimar Wolfenbüttel, 2015), http://nbn-resolving.de/10.17175/sb001_020.

33 Als eines der ältesten Publikationsplattformen und Volltextserver „Art Dok,“ Repositorium für die Kunstgeschichte (zuletzt geprüft am 13.02.2018, <http://archiv.ub.uni-heidelberg.de/artdok/>), seit wenigen Wochen online das „GenderOpen-Repository“ (zuletzt geprüft am 13.02.2018, <https://www.genderopen.de/>), am 1. Oktober 2018 startet „Archivum Medii Aevi Digitale (AMAD),“ ein Repositorium für Open Access-Publikationen aus der Mediävistik, Karoline Döring, „DFG-Projekt „Archivum Medii Aevi Digitale – Mediävistisches Fachrepositorium und Wissenschaftsblog für die Mittelalterforschung (AMAD);“ *Mittelalter. Interdisziplinäre Forschung und Rezeptionsgeschichte* (Blog), 6. Februar 2018, zuletzt geprüft am 19.02.2018, <http://mittelalter.hypotheses.org/12026>.

34 Hubertus Kohle, „Repositorien: vom Datencontainer zum Diskursraum,“ *arthistoricum.net* (Blog), 9. Dezember 2017, zuletzt geprüft am 13.02.2018, <https://blog.arthistoricum.net/beitrag/2017/12/09/repositorien-vom-datencontainer-zum-diskursraum/>.

grob in institutionsweite³⁵ oder übergreifende³⁶ Angebote einordnen. Im ersteren Falle erfolgt konsequenterweise auch der Nachweis über institutionsgebundene Verzeichnisse, etwa an der Universität Heidelberg als Teilkatalog der Hochschulbibliographie, in der Daten zu allen Fachgebieten verzeichnet sind.³⁷

Wie die Forschenden aber für die eigene Arbeit relevante Datensammlungen suchen und finden können, ohne jedes Repositoryum einzeln abzufragen, ist noch offen, denn es fehlt an zentralen Nachweisinstrumenten – es fehlt der Wegweiser in Form domänenspezifischer Kataloge, aber Räume wie die Publikationsrepositorien gibt es auch nicht. Existierende Ansätze sind etwa aggregierende Metalösungen wie re3data³⁸ als Register von Forschungsdatenrepositorien oder OpenAIRE³⁹ als Zugang zu den Ergebnissen aus EU-Projekten. Viele Vorhaben legen (Zwischen-)Ergebnisse und Dokumentation in GitHub⁴⁰ ab, ein (kommerzielles) Repositoryum, das hauptsächlich auf Code ausgerichtet ist, in dem aber auch Textdateien oder Bilder gespeichert werden können – die Auffindbarkeit ist hier ohne Wegweiser auch nicht gegeben. Ein noch im Aufbau befindliches Konzept zur Beschreibung und Verzeichnung ist die DARIAH-DE Collection Registry⁴¹, ein Katalog für Forschungsdatensammlungen. Diese sind mit DCDDM, dem an Dublin Core angelehnten DARIAH Collection Data Description Modell,⁴² standardisiert beschrieben und damit auch maschinell nutzbar.

Die Forschenden wählen das Repositoryum für die Datenpublikation selbst, insofern muss die Verzeichnung und damit die Voraussetzung für die Auffindbarkeit andernorts umgesetzt werden – die Kompetenz und Erfahrung hierfür ist in den Bibliotheken traditionell vorhanden.⁴³ Denn auch Datenpublikationen (z.B. in einem Repository publiziert, mit DOI und ORCID versehen) können katalogisiert werden, analog zu gedruckten Publikationen. Allerdings ist dies bei Daten komplexer, erfordert unter Umständen domänenspezifisches Wissen und ist (noch) ohne institutionelle Strukturen und ohne etablierte Workflows. Auffindbar würden diese Publikationen dann über die einschlägigen Kataloge oder Projekte wie die an der Bielefelder Universitätsbibliothek angesiedelte Suchmaschine BASE.⁴⁴ Doch die Sichtbarkeit und Zugänglichkeit von Forschungsdaten kann auch aktiv von den Forschenden

35 Wie z.B. das „Servicezentrum Forschungsdaten und -publikationen“ an der TU Berlin (Kooperation von UB, Rechenzentrum und Forschungsabteilung), zuletzt geprüft am 13.02.2018, www.szf.tu-berlin.de oder „heiDATA“, das Repositoryum des Competence Centre for Research Data der Universität Heidelberg, zuletzt geprüft am 13.02.2018, <https://heidata.uni-heidelberg.de/>.

36 Für die Geisteswissenschaften „DARIAH-DE Repository“, zuletzt geprüft am 18.02.2018, <https://de.dariah.eu/repository>; vgl. auch „User Guide, Das DARIAH-DE Repository und das TextGrid Repository“, vom 14.02.2018, zuletzt geprüft am 18.02.2018, <https://wiki.de.dariah.eu/display/publicde/Das+DARIAH-DE+Repository+und+das+TextGrid+Repository>; „CLARIN-D Repository“, zuletzt geprüft am 19.02.2018, <http://www.sfs.uni-tuebingen.de/ascl/clarin-center/repository.html>.

37 „Heidelberger Forschungsdatenverzeichnis“, zuletzt geprüft am 19.02.2018, <http://data.uni-heidelberg.de/datenverzeichnis.html>.

38 „Registry of Research Data Repositories“, zuletzt geprüft am 19.02.2018, <https://www.re3data.org/>.

39 „OpenAIRE“, zuletzt geprüft am 19.02.2018, <https://www.openaire.eu/>.

40 „GitHub“, zuletzt geprüft am 19.02.2018, <https://github.com/>.

41 „Informationen zur Collection Registry“, zuletzt geprüft am 19.02.2018, <https://de.dariah.eu/collection-registry>; „Collection Registry“, zuletzt geprüft am 19.02.2018, <https://colreg.de.dariah.eu/colreg-ui/>.

42 „DCDDM“, zuletzt geprüft am 19.02.2018, <https://github.com/DARIAH-DE/DCDDM>.

43 Roland Bertelmann und Hans Pfeiffenberger, „Forschungsdaten und Bibliotheken“, in *Praxishandbuch Bibliotheksmanagement*, hrsg. Rolf Griebel, Hildegard Schäffler und Konstanze Söllner (Berlin, Boston: De Gruyter Saur, 2014), 651-664.

44 „Bielefeld Academic Search Engine“, zuletzt geprüft am 19.02.2018, <https://www.base-search.net/>.

den selbst erhöht werden: Anders als in den Naturwissenschaften werden die Daten in der Literatur bisher häufig nicht zitiert, weshalb auch keine Notwendigkeit zu bestehen scheint, überhaupt welche zu publizieren.

6. Kulturwandel vor Technologie

Die technologischen Voraussetzungen wie Repositorien, Schnittstellen, Standardisierung sind weitgehend entwickelt. Die Technologie stellt damit nicht mehr die wesentliche Barriere dar.⁴⁵ Ob die Publikation von Forschungsdaten zu einer gelebten wissenschaftlichen Praxis wird, ist eng an Veränderungen des Wissenschaftssystems geknüpft. Die Grundvoraussetzung für das Teilen von Forschungsdaten ist die Bereitschaft der Forschenden. Dies erfordert jedoch einen Kulturwandel der zentralen Akteure des Wissenschaftsbetriebes. Während Technologien und Methoden weit fortgeschritten sind, fehlen Abläufe, Strukturen, Tradition, Kreditierung und nicht selten ein grundsätzliches Bewusstsein für die Relevanz von Forschungsdaten.⁴⁶ Diese Faktoren gelten fast ausnahmslos für alle Wissenschaften, aber für einen Kulturwandel müssen die gerade die disziplinspezifischen Eigenheiten berücksichtigt werden:

1. Wie hier ausgeführt wird, hat sich in den Fächern noch keine gemeinsame Begriffsdefinition für geisteswissenschaftliche Forschungsdaten herausgebildet und die datengestützte Forschung in Geisteswissenschaften ist selbst noch in der Entwicklungsphase. Das hat Auswirkungen auf das Reputationssystem, bei Evaluationen und Berufungen spielen Forschungsdaten keine Rolle.
2. Das traditionelle Verständnis wissenschaftlicher Nachvollziehbarkeit der Forschungsergebnisse basiert auf der Identifizierung der Quellen durch Verzeichnisse (Fußnotenapparat, Quellen- und Literaturverzeichnis), nicht durch die unmittelbare Bereitstellung der Quellen. Eine Nachnutzbarkeit ist (noch) nicht intendiert.⁴⁷
3. Anders als in naturwissenschaftlichen Feldern erhalten viele Forschungsdaten ihren Wert als Untersuchungsgegenstand über längere Zeit. Nicht die erste Entdeckung und Analyse, sondern die beste Interpretation wird als höchstes gewichtet.

Die Diskussion um die Publikation von Forschungsdaten wird vor allem von den Naturwissenschaften vorangetrieben. Spezifisch geisteswissenschaftliche Herausforderungen bleiben so häufig unberücksichtigt. Hier lassen sich Parallelen zur Open-Access-Bewegung identifizieren und darin liegt eine große Gefahr. In den Geisteswissenschaften finden sich viele Stimmen, die Open Access als

45 Obwohl es noch Bedarf an technologischer Entwicklung gibt, konstatiert daher eine Expertengruppe für die European Open Science Cloud: „The majority of the challenges to reach a functional EOSC are social rather than technical.“ The Commission High Level Expert Group on the European Open Science Cloud (EOSC), *Realising the European Open Science Cloud* (Luxembourg: Publications Office of the European Union, 2016), 6, <http://dx.doi.org/10.2777/940154>.

46 Eine gute Übersicht über die Barrieren in allen Wissenschaften liefert Christine Borgman, *Scholarship in the digital age: information, infrastructure, and the Internet* (Cambridge, London: MIT Press, 2007).

47 Siehe hierzu Andorfer, *Forschungsdaten*, 19.

übergestülptes und nicht passfähiges Konzept aus den Naturwissenschaften zurückweisen.⁴⁸ Die geisteswissenschaftlichen Fächer müssen daher eigene Konzepte für eine Publikationspraxis von Forschungsdaten entwickeln. Diese Konzeption steht vor der Herausforderung, die fachspezifischen Bedürfnisse insoweit zu berücksichtigen, dass sie mit den etablierten technologischen und informationswissenschaftlichen Standards umzusetzen sind. Idealerweise erfolgt dies gemeinsam mit den Informationsinfrastruktureinrichtungen, die die Forschungsdaten nach der Publikation bereitstellen und betreuen.

7. Kuration statt Friedhof

Die DFG empfiehlt als gute wissenschaftliche Praxis für die einer Publikation zugrunde liegenden Daten eine „Langzeitarchivierung“ von zehn Jahren.⁴⁹ Geisteswissenschaftliche Forschungsdaten sind tendenziell aber viel länger interessant oder bekommen, vom momentanen Standpunkt aus vielleicht gar nicht vorhersehbar, erst viel später eine erneute Relevanz. Dabei steht weniger die Reproduzierbarkeit der Ergebnisse im Vordergrund, sondern vielmehr die Nachnutzung der gesammelten (und bearbeiteten) Quellen durch andere Forschende sowie die langfristige Sicherung von hochwertigen Digitalisaten, Volltexten etc.

Bei der Erhaltung der Daten und dem Zugang zu ihnen sind verschiedene Aspekte zu unterscheiden, die alle ihrem jeweils spezifischen Alterungsprozess unterliegen und die Daten zu unterschiedlichen Zeitpunkten „sterben“ lassen können: Die (technische) Aktualität der Inhalte, der Erschließung und schließlich des technischen Systems. Um einem Datenfriedhof rechtzeitig entgegenzuwirken, ist die Kuration von Forschungsdaten zur Sicherstellung ihrer Unversehrtheit, Lesbarkeit und Interpretierbarkeit unabdingbar.⁵⁰

Da in den Geisteswissenschaften häufig projektbezogen, also mit begrenzter Zeit und limitierten Ressourcen geforscht wird, die Daten aber über das Projektende hinaus aufbewahrt und gepflegt werden müssen, entstehen sog. Humanities Data Center, die die Archivierung, Bereitstellung und Präsentation dieser Daten langfristig begleiten und übernehmen, wobei die Organisation und Zuständigkeit dieser Datenzentren sehr divers ausgestaltet ist.⁵¹

48 Übersichtlich referiert im Debattenbeitrag von Ben Kaden, „Publikationsfreiheit.de, Open Access und Geisteswissenschaften,“ *LIBREAS* (Blog), 31. Juli.2017, zuletzt geprüft am 13.02.2018, https://libreas.wordpress.com/2017/07/31/open-access_publikationsfreiheit/.

49 Deutsche Forschungsgemeinschaft (DFG), *Sicherung Guter Wissenschaftlicher Praxis*, Weinheim: Wiley-VCH, 2013, <http://dx.doi.org/10.1002/9783527679188.oth1>.

50 Michael Lautenschlager, „Institutionalisierte ‚Data Curation Services‘,“ in *Handbuch Forschungsdatenmanagement*, hrsg. Stephan Büttner, Hans-Christoph Hobohm und Lars Müller (Bad Honnef: BOCK + HERCHEN, 2011), 149, <http://nbn-resolving.de/urn:nbn:de:kobv:525-opus-2412>.

51 DHd AG Datenzentren, *Geisteswissenschaftliche Datenzentren im deutschsprachigen Raum. Grundsatzpapier zur Sicherung der langfristigen Verfügbarkeit von Forschungsdaten* (Hamburg: DHd e.V., 2017), <http://dx.doi.org/10.5281/zenodo.1134759>.

8. Lizenzen für Gesetzeshürden

Eine restriktive und komplexe Rechtslage kann für geisteswissenschaftliche Forschungsdaten als charakteristisch bezeichnet werden. Eine verlässliche Rechtslage ist für die Forschenden und publizierenden Einrichtungen jedoch ein entscheidender Faktor, der damit die Zugänglichkeit und die Verwertbarkeit der Forschungsdaten unmittelbar beeinflusst. Zum einen stehen Untersuchungsgegenstände wie zeitgenössische Werke unter urheberrechtlichem Schutz. Aber auch wenn ein Kulturgut selbst nicht mehr dem Urheberrecht unterliegt, kann dessen digitale Reproduktion geschützt und damit die Publikation dieser Forschungsdaten ohne Einigung mit dem Rechteinhaber nicht möglich sein. Der Gegenstand der Forschung kann auch weitere Rechtsbereiche berühren, etwa Persönlichkeitsrechte oder Datenschutz.⁵² Noch komplexer wird die Situation, wenn es sich – wie häufig in den Geisteswissenschaften – um eine Sammlung handelt, sei es eine Zusammenstellung ähnlicher Inhalte, beispielsweise mehrerer Handschriftendigitalisate, oder eine Sammlung im weiteren Sinne, die unterschiedliche Daten wie Digitalisate, Metadaten, eine Datenbank und Auswertungen enthält.⁵³

Die Lösung gegen unklare Rechtslage und Unsicherheit wäre, eine – wo möglich – freie Lizenz wie eine der Creative Commons-Lizenztypen⁵⁴ zu vergeben.⁵⁵ Somit wissen Forschende auf einen Blick, was sie unter welchen Bedingungen mit den Daten machen können und ob diese veröffentlicht werden dürfen. Dabei kommt es darauf an, was im Einzelnen lizenziert wird, ob die Lizenz die Daten selbst, die Datenbank an sich oder sogar die Metadaten und Beschreibung der Sammlung betrifft.

Ein bislang noch ungelöstes Problem sind Sperrfristen, die nach einer bestimmten Zeit ablaufen oder durch neue rechtliche Bedingungen verkürzt werden könnten.⁵⁶ Theoretisch könnte dann die Lizenz angepasst, der Zugang ermöglicht werden, doch hierfür gibt es keinen Automatismus, auch hier muss eine Form von Data Curation betrieben werden.

- 52 Auf die rechtlichen Bestimmungen kann an dieser Stelle nicht im Detail eingegangen werden. Einstiegspunkte für eine weitere Beschäftigung im Bereich der Geisteswissenschaften wären beispielsweise: Paul Klimpel und John H. Weitzmann, *Forschen in der digitalen Welt: Juristische Handreichung für die Geisteswissenschaften*. DARIAH-DE Working Papers Nr. 12 (Göttingen: DARIAH-DE, 2015), <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-5-0>; Nikolaos Beer et al., *Datenlizenzen für geisteswissenschaftliche Forschungsdaten. Rechtliche Bedingungen und Handlungsbedarf*. DARIAH-DE Working Papers Nr. 6 (Göttingen: DARIAH-DE, 2014), <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2014-4-8>; „Portal zu Forschungslizenzen,“ zuletzt geprüft am 16.02.2018, <http://forschungslizenzen.de/>.
- 53 Vgl. zu einer inklusiven Definition von „Forschungsdaten“ und „wissenschaftlicher Sammlung“ Jenny Oltersdorf und Stefan Schmunk, „Von Forschungsdaten und wissenschaftlichen Sammlungen. Zur Arbeit des Stakeholdergremiums ‚Wissenschaftliche Sammlungen‘ in DARIAH-DE,“ *Bibliothek Forschung und Praxis* 40, Nr. 2 (2016): 181f., <http://dx.doi.org/10.1515/bfp-2016-0036>.
- 54 „Erläuterung zu Creative Commons,“ zuletzt geprüft am 16.02.2018, <https://de.creativecommons.org/index.php/was-ist-cc/>.
- 55 „Übersicht über freie und restriktive Lizenzmodelle,“ zuletzt geprüft am 16.02.2018, <http://www.ifross.org/lizenz-center/>.
- 56 Eric Steinhauer, „Recht als Risiko für das kulturelle Erbe,“ in *Mit gutem Recht erinnern*, hrsg. Paul Klimpel (Hamburg: Hamburg University Press, 2018), 134-136, <http://dx.doi.org/10.15460/HUP.178>.

9. Dilemma mit Potenzial

Sind die geschilderten Faktoren wie die Integration der Forschungsdaten in ein vertrauenswürdiges Repositorium, der Eintrag in einem wissenschaftlich relevanten Nachweissystem, die Vergabe einer freien Lizenz und die Anlage einer Dokumentation für eine Forschungsdatensammlung erfüllt, so stehen sie der Wissenschaft für die Nachnutzung in einer adäquaten Form zur Verfügung. Im idealen Fall werden die Forschungsdaten direkt mit der Publikation verbunden: So sind in einem bildanalytischen Aufsatz von Waltraud von Pippich die Messwerte und Metadaten, die der Untersuchung zugrunde liegen, über Links in das Forschungsdatenrepositorium der Universität München nachgewiesen.⁵⁷ Das Beispiel demonstriert auch die Wichtigkeit der Persistent Identifier für die wissenschaftliche Verwendung von Forschungsdaten und die Relevanz von nachhaltiger Infrastruktur.⁵⁸

Es ist offensichtlich, dass die Publikation von Forschungsdaten an inhaltliche und technische Kriterien geknüpft ist, die zu einem erhöhten Arbeitsaufwand für die Forschenden führen. Der Mehraufwand kann nur bedingt durch andere technische oder manuelle Unterstützung gemindert werden, da verschiedene Prozesse an das Fachwissen des/der Domänenspezialist/inn/en gebunden sind. Nichtsdestotrotz sollten die Forschenden sich nicht noch in neue technische Kontexte einarbeiten müssen, die für die Publikation der Forschungsdaten relevant sind. Gerade in diesem Kontext sollten wissenschaftliche Bibliotheken ein Angebot für ihre Community bereitstellen, welches zumindest die Beratung umfasst, falls keine eigene Infrastruktur zur Verfügung steht. Es ist auch vorstellbar, dass die Fachinformationsdienste hier verstärkt aktiv werden. Ein *do-ut-des*-Szenario würde in der Unterstützung der Aufbereitung, Publikation und langfristigen Verfügbarkeit der Forschungsdaten auf der einen und in dem Einbringen von Zeit und Fachwissen auf der anderen Seite liegen.

Bevor sich ein solches Konzept etablieren kann, müssen „neue Blicke durch die alten Löcher“⁵⁹ geworfen werden, um es mit Georg Christoph Lichtenberg zu sagen. Denn trotz der Heterogenität der Forschungsdaten in den Geisteswissenschaften handelt es sich bei vielen der in diesem Beitrag skizzierten Fragen und Aspekte nicht um unüberwindbare Hindernisse, die außerhalb eines paradigmatischen Forschungsprozesses liegen. Das Dilemma besteht in der Frage nach dem Verhältnis von Relevanz, Investition und Ertrag – sprich dem Bewusstsein für die Bedeutung von Forschungsdaten, dem Mehraufwand der Publikation und deren wissenschaftlicher Reputation. Diese Misere kann nicht einseitig gelöst werden, aber an den Lehrstühlen können Praktiken vorgelebt und vermittelt werden, an den Universitäten können Leistungen anerkannt und honoriert werden und bei den Fördergebern die Bereitstellung und Nachnutzung von Forschungsdaten in den Geisteswissenschaften gefördert

57 Vgl. Waltraud von Pippich, „Rot rechnen,“ in *Grenzen und Möglichkeiten der Digital Humanities*, hrsg. Constanze Baum und Thomas Stäcker. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1 (Wolfenbüttel: Herzog August Bibliothek und Forschungsverbund Marbach Weimar Wolfenbüttel, 2015), http://dx.doi.org/10.17175/sb001_016. Zu den Forschungsdaten vgl. Waltraud von Pippich, *Frequenzen und statistische Dispersion der Farben in 50 Bildern von Adolph Menzel* (München: Open Data LMU, 2014), <http://dx.doi.org/10.5282/ubm/data.79>.

58 Vgl. Kathrin Schroeder, „Persistent Identifier (PI) – ein Überblick,“ in *Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*, hrsg. Heike Neuroth et al., Version 2.3 (Göttingen: nestor, 2010), Kap. 9.4, 9:22-45, <http://nbn-resolving.de/urn:nbn:de:0008-2010071949>. Zu den infrastrukturellen Kriterien vgl. das DINI-Zertifikat, zuletzt geprüft am 23.01.2018, <https://dini.de/dini-zertifikat/>.

59 Georg Christoph Lichtenberg, *Sudelbuch F, 1776-1779, F 879*, zit. nach „Projekt Gutenberg-DE,“ zuletzt geprüft am 19.02.2018, <http://gutenberg.spiegel.de/buch/-6445/7>.

und gefordert werden. Geisteswissenschaftliche Infrastrukturen schließlich können die Voraussetzungen dafür schaffen, dass Forschende nicht noch eine technische *terra incognita* betreten müssen, wenn sie ihre Daten publizieren wollen.

Das Unterstützungsangebot der Informationsinfrastruktureinrichtungen nimmt dabei eine Sonderrolle ein, es muss bereits etabliert sein, wenn sich die anderen Akteure bewegen. Erst dann kann auch die Datenautorschaft die Rolle einnehmen, die ihr in einer zunehmend digital geprägten Forschungskultur gebührt. Denn Forschungsdaten sind eben kein von den Naturwissenschaften geborgter Begriff, sondern vielmehr essentieller Bestandteil einer auf digitalen Formaten basierenden geisteswissenschaftlichen Forschungslandschaft.

Abbildungsnachweis:

- Abb. 1: „Titelblatt des Auktionskatalog von Benedikt Bahnsen,“ Digitalisat der Herzog August Bibliothek Wolfenbüttel, zuletzt geprüft am 23.01.2018, <http://diglib.hab.de/drucke/bc-kapsel-7-23s/start.htm?image=00001>.
- Abb 2: „Visualisierung der Druckorte / Druckjahre des Auktionskataloges von Benedikt Bahnsen im DARIAH-DE Geobrowser,“ zuletzt geprüft am 01.02.2018, <https://geobrowser.de.dariah.eu/?csv1=http://geobrowser.de.dariah.eu./storage/267851¤tStatus=mapChanged=Historical+map+of+1650>.
- Abb. 3: „Mischwesen. Datenbankeintrag zu Mischwesen,“ (TA 1680, Tafel CC, 6), aus: Joachim von Sandrart. *Teutsche Academie der Bau-, Bild- und Mahlerey-Künste, Nürnberg 1675-1680, wissenschaftlich kommentierte Online-Edition*, hrsg. von Th. Kirchner et al., 2008-012, zuletzt geprüft am 19.02.2018, <http://ta.sandrart.net>.

Literaturverzeichnis

- Andorfer, Peter. *Forschungsdaten in den (digitalen) Geisteswissenschaften: Versuch einer Konkretisierung*. DARIAH-DE Working Papers 14. Göttingen: DARIAH-DE, 2015. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-7-2>.
- Baum, Constanze und Thomas Stäcker. „Die Digital Humanities im deutschsprachigen Raum. Methoden – Theorien – Projekte.“ In *Grenzen und Möglichkeiten der Digital Humanities*, herausgegeben von Constanze Baum und Thomas Stäcker. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1. Wolfenbüttel: Herzog August Bibliothek und Forschungsverbund Marbach Weimar Wolfenbüttel, 2015. http://dx.doi.org/10.17175/sb001_023.
- Beer, Nikolaos, Kristin Herold, Maurice Heinrich, Wibke Kolbmann, Thomas Kollatz, Matteo Romanello, Sebastian Rose, Felix Falco Schäfer und Niels-Oliver Walkowski. *Datenlizenzen für geisteswissenschaftliche Forschungsdaten. Rechtliche Bedingungen und*

Handlungsbedarf. DARIAH-DE Working Papers Nr. 6. Göttingen: DARIAH-DE, 2014. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2014-4-8>.

- Bertelmann, Roland und Hans Pfeiffenberger. „Forschungsdaten und Bibliotheken.“ In *Praxishandbuch Bibliotheksmanagement*, herausgegeben von Rolf Griebel, Hildegard Schäffler und Konstanze Söllner, 651-664. Berlin, Boston: De Gruyter Saur, 2014.
- Beyer, Hartmut, Jörn Münkner, Katrin Schmidt und Timo Steyer. „Bibliotheken im Buch: Die Erschließung von privaten Büchersammlungen der Frühneuzeit über Auktionskataloge.“ In *Kodikologie und Paläographie im digitalen Zeitalter 4 (Codicology and Palaeography in the Digital Age)*, herausgegeben von Hannah Busch, Franz Fischer und Patrick Sahle, 43-70. Norderstedt: Books on Demand, 2017. <http://nbn-resolving.de/urn:nbn:de:hbz:38-77794>.
- Borgman, Christine. *Scholarship in the digital age: information, infrastructure, and the Internet*. Cambridge, London: MIT Press, 2007.
- Buddenbohm, Stefan, Claudia Engelhardt und Ulrike Wuttke. „Angebotsgenese für ein geisteswissenschaftliches Forschungsdatenzentrum.“ *Zeitschrift für digitale Geisteswissenschaften* (2017). http://dx.doi.org/10.17175/2016_003.
- EOSC, The Commission High Level Expert Group on the European Open Science Cloud. *Realising the European Open Science Cloud*. Luxembourg: Publications Office of the European Union, 2016. <http://dx.doi.org/10.2777/940154>.
- Degkwitz, Andreas. „Enhanced Publications Exploit the Potential of Digital Media.“ In *Evolving Genres of ETDs for Knowledge Discovery. Proceedings of ETD 2015*, 18th International Symposium on Electronic Theses and Dissertations. 51-59. Zuletzt geprüft am 19.02.2018, https://www.researchgate.net/profile/Andreas_Degkwitz/publication/283783359_Enhanced_Publications_Exploit_the_Potential_of_Digital_Media/links/5646fe2908ae9f9c13e8c4fa/Enhanced-Publications-Exploit-the-Potential-of-Digital-Media.pdf.
- DFG, Deutsche Forschungsgemeinschaft. *Sicherung Guter Wissenschaftlicher Praxis*. Weinheim: Wiley-VCH, 2013. <http://dx.doi.org/10.1002/9783527679188.oth1>.
- DHd AG Datenzentren. *Geisteswissenschaftliche Datenzentren im deutschsprachigen Raum. Grundsatzpapier zur Sicherung der langfristigen Verfügbarkeit von Forschungsdaten*. Hamburg: DHd e.V., 2017. <http://dx.doi.org/10.5281/zenodo.1134759>.
- Döring, Karoline. „DFG-Projekt ‚Archivum Medii Aevi Digitale – Mediävistisches Fachrepositorium und Wissenschaftsblog für die Mittelalterforschung (AMAD)‘.“ *Mittelalter. Interdisziplinäre Forschung und Rezeptionsgeschichte* (Blog). 6. Februar 2018. Zuletzt geprüft am 19.02.2018. <http://mittelalter.hypotheses.org/12026>.

- Federbusch, Maria. „OCR für Drucke der Frühen Neuzeit? Erfahrungen und Perspektiven am Beispiel von Funeralschriften.“ *Bibliotheksdienst* 49, Nr. 7 (2015): 713-724.
- Fiedler, Maik und Andreas Weiß. „Von Daten zu Erkenntnissen: Digitale Geisteswissenschaften als Mittler zwischen Information und Interpretation. DHd-Jahrestagung 2015.“ *H-Soz-Kult* (Blog). 6. Juli 2015. Zuletzt geprüft am 13.02.2018. <http://www.hsozkult.de/conferencereport/id/tagungsberichte-6059>.
- Gradl, Tobias, Andreas Henrich und Christoph Plutte. „Heterogene Daten in den Digital Humanities: Eine Architektur zur forschungsorientierten Föderation von Kollektionen.“ In *Grenzen und Möglichkeiten der Digital Humanities*, herausgegeben von Constanze Baum und Thomas Stäcker. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1. Wolfenbüttel: Herzog August Bibliothek, 2015. http://nbn-resolving.de/10.17175/sb001_020.
- Kaden, Ben. „Publikationsfreiheit.de, Open Access und Geisteswissenschaften.“ *LIBREAS* (Blog), 31. Juli.2017. Zuletzt geprüft am 13.02.2018, https://libreas.wordpress.com/2017/07/31/open-access_publikationsfreiheit/.
- Klaffki, Lisa, Stefan Schmunk und Thomas Stäcker. *Stand der Kulturgutdigitalisierung in Deutschland: Eine Analyse und Handlungsvorschläge des DARIAH-DE Stakeholdergremiums ‚Wissenschaftliche Sammlungen‘*. DARIAH-DE Working Papers Nr. 26. Göttingen: DARIAH-DE, 2018. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2018-1-3>.
- Klaffki, Lisa. „Memorandum: Digitalisierung von Kulturgut.“ *DHd-Blog*, 17. Mai 2018. Zuletzt geprüft am 05.06.2018, <https://dhd-blog.org/?p=9849>.
- Klimpel, Paul und John H. Weitzmann. *Forschen in der digitalen Welt. Juristische Handreichung für die Geisteswissenschaften*. DARIAH-DE Working Papers Nr. 12. Göttingen: DARIAH-DE, 2015. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-5-0>.
- Klump, Jens. „Digitale Forschungsdaten.“ In *Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*, herausgegeben von Heike Neuroth, Karsten Huth, Achim Oßwald, Regine Scheffel und Stefan Strathmann. Version 2.3, Kap. 17.10, 17:104-115. Göttingen: nestor, 2010. <http://nbn-resolving.de/urn:nbn:de:0008-2010071949>.
- Kohle, Hubertus. „Digitales Publizieren.“ In *Digital Humanities: Eine Einführung*, herausgegeben von Fotis Jannidis, Hubertus Kohle und Malte Rehbein, 199-205. Stuttgart: J. B. Metzler, 2017.
- Kohle, Hubertus. „Repositorien: vom Datencontainer zum Diskursraum.“ *arthistoricum.net* (Blog), 9. Dezember 2017. Zuletzt geprüft am 13.2.2018. <https://blog.arthistoricum.net/beitrag/2017/12/09/repositorien-vom-datencontainer-zum-diskursraum/>.

- Knoche, Michael. *Die Idee der Bibliothek und ihre Zukunft*. Göttingen: Wallstein, 2017.
- Kraft, Angelina. „Die FAIR Data Prinzipien für Forschungsdaten.“ *TIB-Blog – Weblog der Technischen Informationsbibliothek (TIB)*, 12. September 2017. Zuletzt geprüft am 24.01.2018. <https://blogs.tib.eu/wp/tib/2017/09/12/die-fair-data-prinzipien-fuer-forschungsdaten/>.
- Lautenschlager, Michael. „Institutionalisierte ‚Data Curation Services‘.“ In *Handbuch Forschungsdatenmanagement*, herausgegeben von Stephan Büttner, Hans-Christoph Hobohm und Lars Müller, 149-156. Bad Honnef: BOCK + HERCHEN, 2011. <http://nbn-resolving.de/urn:nbn:de:kobv:525-opus-2412>.
- Lichtenberg, Georg Christoph. *Sudelbuch F, 1776–1779. F 879*. Zitiert nach „Projekt Gutenberg-DE.“ Zuletzt geprüft am 19.02.2018. <http://gutenberg.spiegel.de/buch/-6445/7>.
- Maier, Petra. *Digital Humanities und Bibliothek als Kooperationspartner*. DARIAH-DE Working Papers Nr. 19. Göttingen: DARIAH-DE, 2016. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2016-5-6>.
- Oltersdorf, Jenny und Stefan Schmunk. „Von Forschungsdaten und wissenschaftlichen Sammlungen: Zur Arbeit des Stakeholdergremiums ‚Wissenschaftliche Sammlungen‘ in DARIAH-DE.“ *Bibliothek Forschung und Praxis* 40, Nr. 2 (2016): 179-185. <http://dx.doi.org/10.1515/bfp-2016-0036>.
- Palmer, Carole L. „Thematic Research Collections.“ In *A Companion to Digital Humanities*, herausgegeben von Susan Schreibman, Ray Siemens und John Unsworth. Kap. 24. Oxford: Blackwell 2004. Zuletzt geprüft am 19.02.2018, <http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-4-5&toc.depth=1&toc.id=ss1-4-5&brand=default>.
- Pompe, Wolfgang. „Geisteswissenschaften.“ In *Langzeitarchivierung von Forschungsdaten: Eine Bestandsaufnahme*, herausgegeben von Heike Neuroth, Stefan Strathmann, Achim Oßwald, Regine Scheffel, Jens Klump, Jens Ludwig, 137-159. Göttingen: Universitätsverlag Göttingen / Boizenburg: Verlag Werner Hülsbusch, 2012. <http://nbn-resolving.de/urn:nbn:de:0008-2012031401>.
- Pippich, Waltraud von. *Frequenzen und statistische Dispersion der Farben in 50 Bildern von Adolph Menzel*. München: Open Data LMU, 2014. <http://dx.doi.org/10.5282/ubm/data.79>.
- Pippich, Waltraud von. „Rot rechnen.“ In *Grenzen und Möglichkeiten der Digital Humanities*, herausgegeben von Constanze Baum und Thomas Stäcker. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1. Wolfenbüttel: Herzog August Bibliothek und

- Forschungsverbund Marbach Weimar Wolfenbüttel, 2015. http://dx.doi.org/10.17175/sb001_016.
- Puhl, Johanna, Peter Andorfer, Mareike Höckendorff, Stefan Schmunk, Juliane Stiller und Klaus Thoden. *Diskussion und Definition eines Research Data LifeCycle für die digitalen Geisteswissenschaften*. DARIAH-DE Working Papers Nr. 11. Göttingen: DARIAH-DE, 2015. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-4-4>.
 - Sahle, Patrick. „Review von ‚Sandart.net‘.“ *RIDE* 1 (2014): Abs. 11. <https://dx.doi.org/10.18716/ride.a.1.5>.
 - Sahle, Patrick und Simone Kronenwett. „Jenseits der Daten: Überlegungen zu Datenzentren für die Geisteswissenschaften am Beispiel des Kölner ‚Data Center for the Humanities‘.“ *LIBREAS. Library Ideas* 23 (2013): 76-96. <http://nbn-resolving.de/urn:nbn:de:kobv:11-100212726>.
 - Sandart, Joachim von. *L'Academia Todesca. della Architectura, Scultura & Pittura: Oder Teutsche Academie der Edlen Bau- Bild- und Mahlerey-Künste. Bd. 1,1*. Nürnberg: 1675. Deutsches Textarchiv, <http://nbn-resolving.de/urn:nbn:de:kobv:b4-30378-3>.
 - Sandart, Joachim von. *Teutsche Academie der Bau-, Bild- und Mahlerey-Künste, Nürnberg 1675–1680, wissenschaftlich kommentierte Online-Edition*, herausgegeben von Thomas Kirchner, Alessandro Nova, Carsten Blüm, Anna Schreurs und Thorsten Wübbena, 2008-2012. Zuletzt geprüft am 19.02.2018, <http://ta.sandart.net>.
 - Schöch, Christoph. „Big? Smart? Clean? Messy? Data in the Humanities.“ *Journal of Digital Humanities* 2, Nr. 3 (2013). Zuletzt geprüft am 24.02.2018. <http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/>.
 - Schroeder, Kathrin. „Persistent Identifier (PI) – ein Überblick.“ In *Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*, herausgegeben von Heike Neuroth, Karsten Huth, Achim Oßwald, Regine Scheffel und Stefan Strathmann. Version 2.3, Kap. 9.4, 9:22-45. Göttingen: nestor, 2010. <http://nbn-resolving.de/urn:nbn:de:0008-2010071949>.
 - Stäcker, Thomas. „Erfahrungsbericht Helmstedter Drucke Online an der Herzog August Bibliothek Wolfenbüttel.“ In *Volltext via OCR. Möglichkeiten und Grenzen. Testszenarien zu den Fundamentalschriften der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz*, herausgegeben von Maria Federbusch und Christian Polzin. Beiträge aus der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz 43, 123-135. Berlin: Staatsbibliothek zu Berlin, 2013).
 - Stäcker, Thomas. „Noch einmal: Was sind geisteswissenschaftliche Forschungsdaten?“ *DHd-Blog*, 6. Dezember 2015. Zuletzt geprüft am 16.01.2018. <http://dhd-blog.org/?p=5995>.

- Steinhauer, Eric. „Recht als Risiko für das kulturelle Erbe.“ In *Mit gutem Recht erinnern*. herausgegeben von Paul Klimpel, 125-136. Hamburg: Hamburg University Press, 2018. <http://dx.doi.org/10.15460/HUP.178>.
- Unsworth, John. *Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?* Vortrag, Humanities Computing: Formal Methods, Experimental Practice. King's College London, 13. Mai 2000. Zuletzt geprüft am 19.02.2018. <http://www.iath.virginia.edu/~jmu2m/Kings.5-00/primitives.html>.

Aufsätze

Open-Access-Monografien in deutschsprachigen Universitätsverlagen

Claudia Schober, Universitätsbibliothek Würzburg

Zusammenfassung:

Universitätsverlage fördern die Publikation wissenschaftlicher Monografien, insbesondere Open-Access-Monografien. Dieses Publikationsformat stellt Verlage unter anderem bei der Frage der Ausgestaltung von Open Access und Themen wie Finanzierung, Kund/inn/en und Qualitätssicherung vor Herausforderungen. Die vorliegende Arbeit untersucht mittels einer qualitativen Befragung, welche Lösungsansätze Universitätsverlage im deutschsprachigen Raum bei diesen Themen verfolgen. Es zeigt sich, dass die Lösungsansätze verschieden sind.

Summary:

University presses support monograph publishing, especially as open access monographs. Challenges of this new publishing model arise in different areas, for example with respect to the design of open access, and topics like business models, customers and quality management. By means of a qualitative survey, this study examines the approaches taken by university presses in German-speaking countries to solve these problems. As can be shown, the approaches are various.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S163-180>

Autorenidentifikation: Schober, Claudia: GND 1158505345

Schlagwörter: Open Access; Monografie; Universitätsverlag

1. Wissenschaftliches Publizieren¹

Eine wichtige Ausdrucksform wissenschaftlicher Kommunikation ist die Publikation. In der Literatur werden als wesentliche Funktionen Registrierung, Zertifizierung, Verbreitung und Archivierung genannt.² Neben den Wissenschaftler/inne/n sind Verlage als Vertriebspartner und Organisatoren der Qualitätssicherung sowie Bibliotheken als Distributoren und Archiv wichtige Akteure dieses Prozesses.³ In Zeiten, in denen Bibliotheken und Verlage von einer Zeitschriften- und teilweise auch einer Monografienkrise sprechen, ausgelöst durch hohe Preissteigerungsraten und stagnierende Bibliotheksetats, stellt sich die Frage, in welcher Form eine Publikation ihre Funktionen noch erfüllen kann. Denn zu den Konsequenzen dieser Krise des Publikationswesens zählt u.a., dass eine Art

1 Der Beitrag basiert auf Daten und Auswertungen, die im Rahmen einer Masterarbeit im weiterbildenden Masterstudiengang Bibliotheks- und Informationswissenschaft der HU Berlin erhoben bzw. erstellt wurden.

2 Niels Taubert, „Formale wissenschaftliche Kommunikation,“ in *Forschungsfeld Wissenschaftskommunikation*, hrsg. Heinz Bonfadelli et al. (Wiesbaden: Springer, 2017), 125–39, https://doi.org/10.1007/978-3-658-12898-2_7.

3 Alice Keller, „Publikationskompetenz als neues Tätigkeitsfeld von Bibliotheken,“ *Bibliotheksdienst* 50, Nr. 7 (2016): 661–671, <https://doi.org/10.1515/bd-2016-0080>.

„Versorgungslücke“ entsteht, so dass Bibliotheken Wissenschaftler/inne/n Literatur nicht mehr im erforderlichen Rahmen zur Verfügung stellen können. Dadurch kann die Situation gerade im Hinblick auf die Verbreitung kritisch werden. Als ein möglicher Lösungsansatz für diese Problematik wird Open Access gehandelt. Ziel ist es dabei wissenschaftliche Informationen frei zur Verfügung zu stellen. Vor diesem Hintergrund werden Universitäten und ihre Bibliotheken selbst aktiv und unterstützen, angehalten beispielsweise durch eine Empfehlung des Wissenschaftsrats, Publikationsaktivitäten z. B. durch die Gründung von Universitätsverlagen.⁴

2. Ziele, Formen und Entwicklung von Universitätsverlagen

Als Ziel oder Auftrag eines Universitätsverlags wird die Verbreitung von wissenschaftlichen Erkenntnissen gesehen, so beispielsweise bei P. Givler.⁵ Sie richten sich dabei gerade an die Autor/inn/en von Publikationen, die es nur schwer in die Programme kommerzieller Verlage schaffen, so beschreibt es S. Thatcher.⁶ Dieses Kriterium findet sich auch in der Richtlinie des Universitätsverlags Göttingen.⁷ Die Gründung von KIT Scientific Publishing, einem der größeren Universitätsverlage im deutschsprachigen Raum, war durch das Ziel, alternative Publikationsformen anzubieten, motiviert, so R. Tobias.⁸ Diese Perspektive auf die Universitätsverlage – als Förderinstrument für alternative Publikationsmodelle – teilen auch der Wissenschaftsrat⁹ und die Deutsche Forschungsgemeinschaft¹⁰, die zur Gründung von Universitätsverlagen aufrufen bzw. diese fördern.

Universitätsverlage werden aber nicht ausschließlich, wie es die Definition nahe legt, von einer Universität als Eigentümerin betrieben.¹¹ Es gibt Universitätsverlage, die den Zusatz „Universitäts-“ lediglich aus historischen Gründen oder wegen indirekter Verbindungen zur Universität führen. Beispiele dafür sind Universitätsverlag Winter¹², Universitätsverlag Röhrig¹³ oder der Leipziger Universitätsverlag¹⁴.

4 Wissenschaftsrat, *Empfehlungen zur digitalen Informationsversorgung durch Hochschulbibliotheken* (Greifswald, 2001), zuletzt geprüft am 21.01.2018, <http://www.wissenschaftsrat.de/download/archiv/4935-01.pdf>.

5 Peter Givler, „University Press Publishing in the United States,“ in *Scholarly Publishing: Books, Journals, Publishers and Libraries in the Twentieth Century*, hrsg. Richard E. Abel und Lyman W. Newman (Wiley, 2002), zuletzt geprüft am 26.04.2018, <http://www.aaupnet.org/about-aaup/about-university-presses/history-of-university-presses>.

6 Sanford G. Thatcher, „From the University Presses: What University Presses Think about Open Access,“ *Against the Grain* 22, Nr. 3 (2010): 56–58, <https://doi.org/10.7771/2380-176X.5570>.

7 Georg-August-Universität Göttingen, *Richtlinie für den Universitätsverlag Göttingen* (09.01.2015), zuletzt geprüft am 21.01.2018, <https://www.univerlag.uni-goettingen.de/dokumente/texts/Nutzungsordnung-2015-01-09.pdf>.

8 Regine Tobias, „Universitätsverlage sind ein wunderbares neues Geschäftsfeld für Bibliotheken,“ *BIT online* 18, Nr. 2 (2015): 170–176, zuletzt geprüft am 21.01.2018, <http://b-i-t-online.de/heft/2015-02-interview-tobias.pdf>.

9 Wissenschaftsrat, *Empfehlungen*.

10 „Aufforderung zur Antragstellung. „Wissenschaftliche Monographien und monographische Serien im Open Access“, Deutsche Forschungsgemeinschaft, zuletzt geprüft am 21.01.2018, http://www.dfg.de/download/pdf/foerderung/programme/lis/ausschreibung_oa_monographien.pdf.

11 Günther Pflug, „Universitätsverlag,“ in *Lexikon des gesamten Buchwesens: Band VII*, hrsg. Severin Corsten et al., 2., völlig neu bearbeitete und erweiterte Aufl. (Stuttgart: Hiersemann, 2007), 596.

12 „Über uns,“ Universitätsverlag Winter GmbH Heidelberg, zuletzt geprüft am 21.01.2018, https://www.winter-verlag.de/de/ueber_uns/.

13 „Verlagsportrait,“ Röhrig Universitätsverlag GmbH, zuletzt geprüft am 21.01.2018, <http://www.roehrig-verlag.de/content/Verlag.html?XTCsid=f38f636954714d968e54d0807a790949>.

14 „Verlagsportrait,“ Röhrig Universitätsverlag GmbH, zuletzt geprüft am 27.06.2018, <https://www.roehrig-verlag.de/über-uns>.

Eine engere Verbindung zur Universität stellen Kooperationen mit kommerziellen Verlagen dar. Beispiel dafür ist der Regensburger Universitätsverlag, bei dem die Universität Regensburg nur einen kleinen Anteil hält. Solche Kooperationen treten zunehmend in Form von Imprint-Modellen auf. Dabei fungiert die Universität als Namensgeber, während der Verlag als Imprintnehmer agiert, der sich Vorteile aus dem Namen der Universität als Label erhofft.¹⁵ Diese Kooperationsform findet sich beispielsweise im Angebot von V&R unipress, bei der ein Großteil der Verlagsarbeit von V&R unipress übernommen wird und die Universität im Rahmen eines wissenschaftlichen Beirats mitwirkt.¹⁶ Es existieren aber auch Modelle, die der Universität mehr Einfluss ermöglichen, wie beispielsweise die unterschiedlichen Angebote für Universitätsverlage von readbox unipress.¹⁷ Eine Variante ist dabei die Organisation einer Schriftenreihe in Kooperation mit readbox unipress, wobei das Institut oder die Hochschule den eigenen ISBN-Nummernkreis nutzt und der Dienstleister neben dem Druck auch den Vertrieb inklusive der Meldung an das Verzeichnis Lieferbarer Bücher und weitere Buchhandelsverzeichnisse übernimmt.

Einige neugegründete Universitätsverlage nutzen solche Imprint-Modelle, z.B. Mainz University Press oder Vienna University Press in Kooperation mit V&R unipress.¹⁸ Daneben gibt es im deutschsprachigen Raum natürlich Universitätsverlage, die im Sinne der Definition von G. Pflug von der Universität als Eigentümerin betrieben werden.¹⁹ Eine Besonderheit des deutschsprachigen Raums ist, dass die Tradition der Universitätsverlage nicht sehr weit zurückreicht. Als ältester Verlag, der von der Universität selbst betrieben wird, gilt der Verlag der Technischen Universität Berlin, gegründet 1969.²⁰ Viele Universitätsverlage wurden erst nach dem Jahr 2000 gegründet.²¹ Dieser Gründungszeitraum der Universitätsverlage im deutschsprachigen Raum ist deutlich kürzer im Vergleich zur langen Tradition der Universitätsverlage beispielsweise im angloamerikanischen Raum, wo die Wurzeln einzelner Verlage bis ins 15. Jahrhundert zurückreichen.²² Dort kam es vor allem in den 1950er Jahren, in denen der Bildungsbereich allgemein stark gefördert wurde, zu einer Hochphase der Verlagsaktivitäten an Universitäten.²³ Charakteristisch für diese Verlage ist die Anbindung an die Universität, die in unterschiedlichem Umfang erfolgen kann, beispielsweise durch ein Reporting gegenüber der Universitätsleitung oder der Bibliothek. Universitätsverlage des deutschsprachigen Raums, die über diese Form der Einbindung verfügen, haben sich z.B. in der Arbeitsgemeinschaft Universitätsverlage

15 Eduard Schönstedt und Thomas Breyer-Mayländer, *Der Buchverlag: Geschichte, Aufbau, Wirtschaftsprinzipien, Kalkulation und Marketing* (Stuttgart: Metzler, 2010), 68–69.

16 Sabine Häußermann, „Aspekte der Gründung eines Universitätsverlages am Beispiel Heidelberg,“ *Berliner Handreichungen zur Bibliotheks- und Informationswissenschaft* 255 (Berlin: Institut für Bibliotheks- und Informationswissenschaft der Humboldt-Universität zu Berlin, 2009), 44, <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:kobv:11-10099919>.

17 „Leistungen,“ readbox unipress, zuletzt geprüft am 21.01.2018, <https://readboxunipress.wordpress.com/leistungen/>.

18 „Verlage der Universitäten Bonn, Mainz, Osnabrück und Wien,“ V&R unipress, zuletzt geprüft am 30.04.2017, <http://www.v-r.de/de/universitaetsverlage/c-377>.

19 Pflug, „Universitätsverlag“.

20 Margo Bargheer und Jutta Pabst, „Being Small is Not a Fault: Making Sense of the Newer Generation of German-Language University Presses,“ *Learned Publishing* 29 (2016): 335–341, <https://doi.org/10.1002/leap.1053>.

21 Ebd.

22 Axel Halle, „Universitätsverlage. Eine vergleichende Perspektive,“ *Zeitschrift für Bibliothekswesen und Bibliographie* 51, Nr. 5/6 (2004): 277–283, zuletzt geprüft am 26.04.2018, http://zfbib.thulb.uni-jena.de/servlets/MCRFileNodeServlet/jportal_derivate_00001709/j04-h5-6-auf-2.pdf.

23 Alison Mudditt, „The Past, Present, and Future of American University Presses: A View from the Left Coast,“ *Learned Publishing* 29 (2016): 330–334, <https://doi.org/10.1002/leap.1047>.

organisiert.²⁴ Die Einbindung bzw. Anbindung an die Universität bedeutet, so M. Bargheer und J. Pabst, dass die Universität oder Bibliothek einen starken Einfluss auf die Entwicklung des Verlages nehmen kann, so auf Budget, Personal, Verlagsprogramm und Zielgruppen.²⁵ Dabei ist Open Access als ein Publikationsmodell natürlich auch ein Thema; die Arbeitsgemeinschaft der Universitätsverlage legt sogar Open-Access-Affinität als Mitgliedschaftskriterium an.²⁶

Wieso besteht aber überhaupt eine Notwendigkeit, Open-Access-Publikationen durch Universitätsverlage zu fördern?

3. Monografien und Open Access

Ein einheitliches Verständnis von Open Access findet man nicht überall. Nicht ohne Grund wird im Mission Statement der Berlin 11 Conference on Open Access die Etablierung eines Standards gefordert.²⁷ Basis dafür und auch für die Definition von Open Access an anderen Stellen ist die Berliner Erklärung.²⁸ In ihrem Sinn bedeutet Open Access, dass einige wesentliche Unterschiede zu herkömmlichen Veröffentlichungswegen bestehen:

- Freier Zugang zur Publikation für den Leser, damit zwingend eine elektronische Veröffentlichung
- Einräumung von Rechten zur Nachnutzung, wie beispielsweise Verbreitung, Bearbeitung, öffentliche Wiedergabe u. v. m. sofern die Urheberschaft korrekt angegeben wird
- Speicherung der Publikation in einem Online-Archiv zur Langzeitarchivierung

Sofern diese Kriterien unmittelbar erfüllt werden, spricht man von Gold Open Access. Dieses Modell der Veröffentlichung ist im STM-Bereich, in dem die Artikelpublikation das wesentliche Kommunikationsmedium ist, bereits gängig. Für den Bereich der Monografien verhält sich dies etwas anders.²⁹ Die Monografie, eine umfassende ausführliche Darstellung zu einem Thema, oft sehr spezialisiert,³⁰ gilt im Bereich der Geistes- und Sozialwissenschaften als eines der wichtigen Kommunikationsmedien. Für diese Form ist die Umsetzung von Open Access aus unterschiedlichen Gründen schwierig. V. Schallehn und R. Schimmer nennen unter anderem die Langlebigkeit der Monografie, z.B. im Hinblick auf die Relevanz der präsentierten Forschungsergebnisse, den im Vergleich zu Zeitschriften langsameren Übergang zu digitalen Publikationsformen und den Lesekomfort der gedruckten

24 „Kriterien für Mitgliedschaft,“ Arbeitsgemeinschaft der Universitätsverlage, zuletzt geprüft am 21.01.2018, https://blog.bibliothek.kit.edu/ag_univerlage/?page_id=912.

25 So zeigt es die Aufzählung, die sich bei Bargheer und Pabst, „Being small is not a fault“ findet.

26 Vgl. Anm. 24.

27 „Mission Statement at the Berlin 11 Open Access Conference of the Max Planck Society (20 Nov 2013),“ zuletzt geprüft am 21.01.2018, https://openaccess.mpg.de/mission-statement_en.

28 „Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities,“ zuletzt geprüft am 21.01.2018, <https://openaccess.mpg.de/Berliner-Erklarung>.

29 Volker Schallehn und Ralf Schimmer, „Open Access,“ in *Praxishandbuch Bibliotheksmanagement*, hrsg. Rolf Griebel et al., Reference (Berlin: de Gruyter Saur, 2016), 311–338, <https://doi.org/10.1515/9783110303261.311>; Frances Pinter, „Open Access for Scholarly Books?,“ *Publishing Research Quarterly* 28, Nr. 3 (2012): 183–191, <http://doi.org/10.1007/s12109-012-9285-0>; Eelco Ferwerda, „Open access monograph business models,“ *Insights: the UKSG journal* 27, s1 (2014): 35–38, <http://doi.org/10.1629/2048-7754.46>.

30 Pinter, „Open Access for Scholarly Books?“.

Ausgabe als Gründe.³¹ Eine Studie von K. Rosenbaum zeigt, dass die Umsetzung von Open Access für Monografien von Fachgebiet zu Fachgebiet sehr unterschiedlich sein kann.³² So spielt Open Access für juristische Publikationen kaum eine Rolle, es wird vorwiegend gedruckt oder in kostenpflichtigen elektronischen Formaten publiziert. Als Ursachen dafür werden bei K. Rosenbaum u.a. die beherrschende Stellung eines Verlags, der große Marktanteil der juristischen Praxis und teilweise bestehende Vergütungsmöglichkeiten z.B. für Kommentare genannt. In der Kunstgeschichte stellt oft die Frage der Bildrechte ein Hindernis dar, da diese zunächst für den entsprechenden Publikationszweck und die jeweilige Publikationsform eingeholt bzw. erworben werden müssen. Dabei werden in der Kunstgeschichte die Vorteile digitaler Publikationen oder paralleler digitaler Publikationen durchaus gesehen. In anderen Fachgebieten, beispielsweise der Wissenschaftsgeschichte, ist Open Access vor allem auch eine Frage der Finanzierung.

Diese Frage nach der Finanzierung ist nicht nur für die Autor/inn/en relevant, sondern auch für die Verlage. Wissenschaftliche Publikationen müssen in vielen Fällen wegen des begrenzten Zielpublikums durch Druckkostenzuschüsse, Subskriptionen oder Ähnliches finanziert werden,³³ so dass der Verlag zumindest keinen Verlust macht.³⁴ Durch den Wegfall von Käufer/inne/n als Bestandteil der Finanzierung verlagert sich diese in vielen Fällen hin zu einem Author-pays-Modell, wie es beispielsweise auch im Zeitschriftenbereich angewendet wird. Die Kosten für die Publikation von Monografien, die ggf. zusätzlich in gedruckter Form erscheinen, sind jedoch deutlich höher als beispielsweise die Kosten einzelner Artikel, die Open Access zur Verfügung gestellt werden.³⁵ Daher werden Alternativen zum Author-pays-Modell gesucht, z. B. in Projekten wie Knowledge Unlatched, das Crowdfunding nutzt.³⁶ Daneben gibt es noch weitere Formen zur Finanzierung u.a. durch die eigene Institution, Förderorganisationen oder Mitgliedschaftsmodelle für Bibliotheken.³⁷ Die Verlagerung der Ausrichtung der Verlage vom bisherigen Fokus auf die Produktion hin zu einem Service

31 Schallehn und Schimmer, „Open Access“.

32 Konstanze Rosenbaum, „Von Fach zu Fach verschieden. Diversität im wissenschaftlichen Publikationssystem,“ in *Wissenschaftliches Publizieren: Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung*, hrsg. Peter Weingart und Niels Taubert (Berlin: de Gruyter, 2016), 41–74, <https://doi.org/10.1515/9783110448115-002>.

33 Niels Taubert, „Open Access und digitale Publikation aus der Perspektive von Wissenschaftsverlagen,“ in Weingart; Taubert, *Wissenschaftliches Publizieren*, 75–102, <https://doi.org/10.1515/9783110448115-003>.

34 Wulf D. von Lucius, *Verlagswirtschaft: ökonomische, rechtliche und organisatorische Grundlagen*, 3., neubearbeitete und erweiterte Aufl. (Konstanz: UVK-Verlagsgesellschaft, 2014), 146–147

35 Article Processing Charges bewegen sich meist in einem Bereich von 1.400 bis 1.600 EUR. Das Projekt OpenAPC, das Angaben zu APC an wissenschaftlichen Institutionen bündelt, verzeichnet in Deutschland einen durchschnittlichen Betrag von 1.417 EUR für Artikel (ohne hybrid) im Jahr 2015 (Bielefeld University). Dabei ist jedoch die Fördergrenze der DFG im Rahmen der Förderlinie „Open Access Publizieren“ zu bedenken, vgl. Johannes Fournier und Roland Weihberg, „Das Förderprogramm »Open Access Publizieren« der Deutschen Forschungsgemeinschaft. Zum Aufbau von Publikationsfonds an wissenschaftlichen Hochschulen in Deutschland,“ *Zeitschrift für Bibliotheks- und Bibliographie* 60, Nr. 5 (2013): 236–243, <https://doi.org/10.3196/186429501360528>. Martin P. Eve, „Open Access publishing and scholarly communications in non-scientific disciplines,“ *Online Information Review* 39, Nr. 5 (2015): 717–732, <https://doi.org/10.1108/OIR-04-2015-0103> gibt für die Publikation von Open-Access-Monografien Beträge von 11.000 GBP bei Palgrave und 15.000 EUR bei Springer an.

36 „Knowledge Unlatched macht weitere 343 Titel Open Access verfügbar,“ Knowledge Unlatched, Pressemitteilung, 23.02.2017, zuletzt geprüft am 21.01.2018, http://content.schweitzer-online.de/static/content/uploads/StoreFront/file/000/110/000110714//Pressemitteilung_KU.pdf.

37 Eelco Ferwerda, Frances Pinter und Niels Stern, *A Landscape Study On Open Access And Monographs: Policies, Funding And Publishing In Eight European Countries* (2017), 36–37, <https://doi.org/10.5281/zenodo.815932>.

für Autor/inn/en hat Konsequenzen für das Handeln der Verlage.³⁸ Denn für die Verlagstätigkeit als Service für Autor/inn/en stellen sich natürlich Fragen nach den verschiedenen Dienstleistungen wie Textsatz oder Korrektorat, ebenso wie die für Autor/inn/en wichtige Frage, wie der Verlag die Sichtbarkeit der Publikationen unterstützt.

Autor/inn/en stellen für jeden Verlag als Inhaltslieferanten eine wichtige Zielgruppe dar. In der Regel erfolgt die Autorenakquisition bei Eigenverlagen, wie es Universitätsverlage sind, eher passiv.³⁹ Bereits C. Woll empfiehlt hier 2010 ein aktiveres Vorgehen.⁴⁰

Gerade bei wissenschaftlichen Publikationen spielt die Qualitätssicherung eine entscheidende Rolle. Dabei geht es nicht nur um die objektiv sichtbare Qualität, sondern vornehmlich um die wissenschaftliche oder inhaltliche Qualität einer Arbeit. Hier ist der Verlag eher als Organisator einer Form von Qualitätssicherung gefragt,⁴¹ beispielsweise durch das Anlegen von Mindestanforderungen bei Prüfungsarbeiten oder durch die Unterstützung bei der Organisation eines Peer-Review-Prozesses oder des Herausbergremiums. Das Peer Review ist ein weiterer Punkt, der den Publikationsprozess von Monografien und Zeitschriften unterscheidet. Die Begutachtung einer Monografie ist langwierig und aufwändig, in vielen buchaffinen Disziplinen kommt die Problematik der Akzeptanz und Etablierung des Peer Review hinzu, da es nicht in allen Fachgebieten gleichermaßen stark verankert ist.⁴²

4. Methode

Die oben geschilderte Problematik, die bei der Publikation von Monografien im Allgemeinen und bei Open-Access-Monografien im Speziellen auftritt, betrifft alle wissenschaftlichen Verlage. Sieht man Universitätsverlage nun als eine Antwort auf die Frage, wie Open Access im Monografienbereich gefördert werden kann, so kann man diesen Gedanken weiter führen zur Frage: Wie gehen Universitätsverlage mit diesen Themen um? Diese Frage lässt sich untergliedern: Wer sind die Autor/inn/en und Kund/inn/en? Welche Dienstleistungen werden geboten? Wie können die Zielgruppen erreicht werden? Wie ist die Qualitätssicherung organisiert? Wie gestaltet sich die Finanzierung? Wie geht der Verlag mit dem Thema Open Access um?

Diese und weitere Fragen wurden im Rahmen einer Online-Befragung im März 2017 an institutionelle Eigenverlage im deutschsprachigen Raum versandt.⁴³ Die Befragung wurde qualitativ gestaltet, um möglichst viele Varianten erfassen zu können. Dies bedeutet für die Ergebnisse, dass alle Angaben

38 Sven Fund, „Will Open Access Change the Game?“, *BIBLIOTHEK Forschung und Praxis* 39, Nr. 2 (2015): 206–209, <https://doi.org/10.1515/bfp-2015-0025>.

39 Verena Huber, „Zur Typologisierung der aktuellen deutschen Verlagslandschaft.“ *Alles Buch : Studien der Erlanger Buchwissenschaft* 46 (Friedrich-Alexander-Universität Erlangen-Nürnberg, 2012), 37, <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:29-opus4-58742>.

40 Christian Woll, *Bibliotheken als Dienstleister im Publikationsprozess: Herausforderungen und Chancen alternativer Formen des wissenschaftlichen Publizierens* (Saarbrücken: VDM Verlag Müller, 2006), 76.

41 Sven Fund, „De Gruyter Open Library – Das Open Access-Modell eines Universalwissenschaftsverlags für Bücher und Zeitschriften.“ *BIBLIOTHEK Forschung und Praxis* 35, Nr. 2 (2011): 157–160, <https://doi.org/10.1515/bfup.2011.021>.

42 Vgl. Rosenbaum, „Von Fach zu Fach verschieden“

43 Der Fragebogen ist im Anhang abgedruckt.

zu Häufigkeiten mit Vorsicht interpretiert werden müssen. Es wurden keine Antwortmöglichkeiten vorgegeben, daher besteht immer die Möglichkeit, dass eine Angabe vielleicht von einem Verlag vergessen oder nicht genannt wurde, da sie z.B. als selbstverständlich oder unwichtig bewertet wurde. Daher sind die Ergebnisse dieser Studie vor allem als Einblick zu verstehen.

Im Mittelpunkt stand dabei die Produktion von Open-Access-Monografien. Um die dabei vorhandene Breite dennoch sinnvoll eingrenzen zu können, wurde der Kreis der angefragten Verlage anhand von drei Kriterien definiert: Zunächst richtete sich die Befragung nur an Verlage des deutschsprachigen Raums, da für sie auch die Problematik eines durch die Publikationssprache unter Umständen zusätzlich eingeschränkten Kundenkreises hinzukommt.⁴⁴ Als weiteres Kriterium wird eine gewisse Professionalisierung des Verlags vorausgesetzt, so dass dieser als solcher nach außen sichtbar ist, und beispielsweise über eine eigene Homepage verfügt. Dabei wurde ein weites Verständnis des Begriffs Universitätsverlag zugrunde gelegt, also der Universitätsverlag im Sinne eines Eigenverlags einer wissenschaftlichen Institution. Als letztes Kriterium wird schließlich die Möglichkeit zur Publikation von Open-Access-Monografien angelegt. Insgesamt konnten so 36 Verlage ermittelt werden,⁴⁵ von denen sich 24 an der Befragung beteiligten.

Die Fragebögen wurden anonymisiert und mit einer qualitativen inhaltlich-strukturierenden Inhaltsanalyse nach U. Kuckartz ausgewertet,⁴⁶ wobei die Kategorien, die den Textstellen zugeordnet wurden, teilweise deduktiv aus den gestellten Fragen, teilweise induktiv aus dem vorliegenden Material generiert wurden. Die übergeordneten Kategorien wie Autorenakquisition, Open Access oder Organisationsform ergaben sich aus dem Fragebogen. Die einzelnen Subkategorien wurden aus den Antworten der Befragten abgeleitet. Beispielsweise für „Teil der Bibliothek“ als Subkategorie zu „Organisationsform“; diese Subkategorie wurde für Antworten wie „Abteilung der Bibliothek“, „Teil der Universitätsbibliothek“, „zur Verfügung gestellt durch die Universitätsbibliothek“ vergeben und somit wurde es möglich die Antworten zu einem gewissen Grad zu bündeln.

5. Ergebnisse

Ausgehend von der gewählten Befragungsmethode können die hier präsentierten Ergebnisse nur als Spiegel dessen, was die Verlage selbst für wichtig erachten, verstanden werden. Die Verlage, die an der Befragung teilgenommen haben, sind in ihrer Zusammensetzung sowohl hinsichtlich des Zeitraums des Bestehens (von über 50 bis 1 Jahr) als auch hinsichtlich der publizierten Bücher (von 1 bis über 200, mit einem Anteil von 0 bis 100 % an Open-Access-Monografien) sehr unterschiedlich.

44 Vgl. Bargheer und Pabst, „Being small is not a fault“

45 Die Zahl von 36 Universitätsverlagen ergab sich aus den Mitgliedern der AG Universitätsverlage (Stand Februar 2017 24, vgl. https://blog.bibliothek.kit.edu/ag_univerlage/?page_id=535) sowie durch Auswertungen des Publisher's International ISBN Directory und einigen persönlich bekannten Verlagen, darunter: Axia Academic Publishers, Edition Donau Universität Krems, Edition Open Access, interact Verlag, Mannheim University Press, Verlag des Römisch-Germanischen Zentralmuseums, Schriftenreihe Schiffsbau, TUM.University Press, Universitätsverlag Graz, Universitätsverlag Hildesheim, Verlag der Fachhochschule Potsdam und der Westdeutsche Universitätsverlag.

46 Udo Kuckartz, *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*, 3., überarbeitete Auflage, Grundlagen-Methoden (Beltz, 2016).

5.1. Autor/inn/en und Autorenakquise

Die Autor/inn/en der befragten Universitätsverlage sind zumeist Wissenschaftler/innen der eigenen Institution, in der Regel beginnend mit Promovierenden, die ihre Dissertation publizieren können. Einzelne Verlage publizieren auch Arbeiten von Masterstudierenden. Dazu kommt die Gruppe „Externer“, die vor allem innerhalb von Schriftenreihen oder Tagungs- oder Sammelbänden die Möglichkeit zur Publikation haben oder durch Kooperationen mit anderen wissenschaftlichen Institutionen vor Ort.

Autor/inn/en gelangen auf unterschiedlichen Wegen zu den Universitätsverlagen. Einzelne Verlage gehen aktiv auf Autor/inn/en zu. Einige bewerben den Verlag im Rahmen von Veranstaltungen, über Social-Media-Kanäle, E-Mails oder setzen auf Herausgeber/innen zur Einwerbung von Autor/inn/en. Die Homepage ist ebenfalls ein häufig genutzter Weg, um Autor/inn/en zu gewinnen. Die Werbung über Mund-zu-Mund-Propaganda wird von vielen Verlagen als Instrument zur Gewinnung von Autor/inn/en bewertet. Auffallend ist, dass die Verlage auf verschiedene Wege zur Gewinnung von Autor/inn/en setzen, nur in wenigen Fällen verlässt sich der Verlag auf einen Weg.

5.2. Dienstleistungen

Die Verlagstätigkeit besteht aus unterschiedlichen Dienstleistungen, die die Verlage für ihre Autor/inn/en erbringen. Zum Kern der Dienstleistungen, die fast alle Verlage nennen, zählen der Bereich Layout und Coverdesign, in den neben dem Textsatz auch das Bereitstellen von Formatvorlagen fällt; daneben auch Druck – in den meisten Fällen wird parallel zur elektronischen Ausgabe eine gedruckte Ausgabe produziert –, in vielen Fällen ist darunter die Kommunikation mit einem Druckdienstleister oder das Einholen von Angeboten zu verstehen, und Vertrieb. In die Kategorie Vertrieb fallen eine Vielzahl einzelner Services, die der Sichtbarkeit der Publikationen dienen, und damit auch in vielen Fällen der Gewinnung von Kund/inn/en. Beispielhaft zu nennen sind hier die Vergabe von DOIs, Eintrag in das Verzeichnis Lieferbarer Bücher, ISBN- und ISSN-Vergabe, Verzeichnung bei Barsortimentern und im Online-Buchhandel, Verteilung im Rahmen des Pflichtexemplarrechts und auch des Schriften-tauschs, Vertrieb über den Online-Shop des Verlags. Um Kund/inn/en oder Leser/innen zu gewinnen, spielen neben der Verzeichnung im Buchhandel und in Bibliothekskatalogen auch die Indexierung in Suchmaschinen wie BASE oder Google Scholar eine Rolle, in einem Fall wird hierfür sogar Suchma-schinenoptimierung betrieben. Auch die auf Open-Access-Monografien spezialisierten Plattformen wie DOAB, OAPEN Library und Open Edition werden von einigen Verlagen genutzt.⁴⁷ Fachlich spe-zialisierte Datenbanken, bei denen sich der Verlag u.U. aktiv um die Aufnahme seiner Publikationen bemüht, werden nur von wenigen Verlagen genannt. Daneben nutzen die Universitätsverlage auch Werbemittel wie Verlagskataloge bzw. -programme, Newsletter, Produktflyer oder Werbepostkarten. Eine physische Bewerbung, so ein gedrucktes Exemplar erstellt wird, wird vereinzelt unter anderem in Form von Vitrinen, auf Messen und Tagungen durchgeführt. Social Media spielt als Werbemittel in diesem Bereich nur eine Nebenrolle. Auch die Cover werden zu Werbezwecken genutzt.

Im Kontext der Verbreitung über den Schriftentausch ist es unter anderem ein Ziel, die Vorgaben der VG Wort, die eine Ausschüttung einer Autorentantieme für die Autoren ermöglichen, zu erfüllen;

47 <https://www.doabooks.org/>; <http://www.oapen.org/>; <https://www.openedition.org/>.

daher wird die VG Wort als weiterer Service genannt.⁴⁸ Nur ein Verlag bietet ein Autorenhonorar ab dem achten verkauften (Druck-)Exemplar an. Vereinzelt nennen die Verlage die Übernahme redaktioneller Aufgaben, die Erstellung von (Nutzungs-)Statistiken und das Einwerben von Drittmitteln. Weitere für die Verlage wichtige Dienstleistungen sind die Autorenbetreuung, also vom Kostenvoranschlag über Fördermöglichkeiten bis zu Fragen der Rechte rund um die Publikation; das Lektorat, dessen Spektrum sehr unterschiedlich gefasst ist – es kann sowohl die Vermittlung von Korrekturen oder Lektoraten oder ein formales Lektorat bedeuten, aber genauso ein Fachlektorat; in Verbindung damit steht die Qualitätssicherung, die explizit als Serviceleistung genannt wurde.

5.3. Qualitätssicherung

Wie funktioniert Qualitätssicherung in den Universitätsverlagen? Für die Beurteilung der Qualität des Produktes Monografie sind sowohl formale als auch inhaltliche Kriterien relevant.

Um die formale Qualität der Bücher, vor allem im Rahmen typografischer Richtlinien, der Druckqualität und der Qualität der Datei, sicherzustellen, werden die meisten Verlage selbst aktiv.

Um inhaltliche Qualität zu überprüfen, lassen sich acht verschiedene Konzepte identifizieren. In einzelnen Fällen kommt eine automatisierte Plagiatsprüfung mittels Software zum Einsatz. Erfahrung mit dem Erscheinungsbild wissenschaftlicher Manuskripte kann ein weiteres Indiz für die inhaltliche Qualität sein. Einige Verlage lassen die Frage der inhaltlichen Qualität vollständig in der Verantwortung der Autor/inn/en. Häufiger vertraut man jedoch auf die Beurteilung des Manuskripts durch einzelne andere Wissenschaftler/innen, beispielsweise im Rahmen von Prüfungsverfahren,⁴⁹ oder durch persönliche Bestätigung. Herausgeber/innen nehmen ebenfalls eine wichtige Rolle bei der Qualitätssicherung ein, insbesondere bei Verlagen, deren Publikationen ausschließlich in Schriftenreihen erscheinen. Einige Verlage arbeiten mit einem Wissenschaftlichen Beirat bzw. einem Editorial Board, das die inhaltliche Prüfung der eingereichten Manuskripte übernimmt. Die Zusammensetzung ist dabei unterschiedlich, wobei in der Regel Vertreter/innen unterschiedlicher Fachdisziplinen im Beirat tätig werden. Eine Prüfung im Rahmen eines Peer-Review-Verfahrens wird nur bei einigen Verlagen angewandt, unter Umständen nur für Publikationen, die durch keinen anderen Prüfungsmechanismus abgedeckt wurden. Nur wenige Verlage verlassen sich allein auf ein Prüfungsverfahren, sondern bringen, teilweise abhängig vom Publikationstyp, unterschiedliche Verfahren zum Einsatz.

5.4. Open Access

Eine zentrale Frage ist, wie die befragten Universitätsverlage das Thema Open Access genau handhaben. Der Anteil der Open-Access-Publikationen in den Verlagsprogrammen variiert deutlich. Ein Teil der Verlage verpflichtet sich ausschließlich Gold Open Access zu publizieren, ein anderer Teil bietet Gold Open Access als eine Publikationsoption neben anderen an. Einige Verlage nutzen den grünen Weg, um Publikationen nach einer Embargofrist von 3 - 12 Monaten zugänglich zu machen. An dieser Stelle weisen einige Verlage ausdrücklich auf die unterschiedlichen fachlichen Gepflogenheiten hin,

48 Dieser Punkt ist vor allem für parallele Druckausgaben relevant. Im Bereich Wissenschaft fordert die VG Wort hier, dass eine angemessene Verbreitung in Bibliotheken erfolgt ist (<http://www.vgwort.de/verguetungen/auszahlungen/wissenschaftliche-publikationen.html>).

49 Ein häufig genanntes Beispiel ist das Erfordernis einer Mindestnote für die Veröffentlichung von Dissertationen.

die in einigen Fällen diese verschiedenen Verfahren bedingen. Neben der Unterscheidung zwischen grünem und goldenem Weg stellt sich die Frage, welche Lizenzen genutzt werden, um den Kriterien, die an Open Access angelegt werden, gerecht zu werden. Dabei finden sich unter den genannten Lizenzen alle Varianten der Creative-Commons-Lizenzen. Einige wenige Verlage legen sich dabei auf die Verwendung einer Lizenz wie CC BY, CC BY-SA oder CC BY-NC-ND fest. An Stelle dieser international bekannten Lizenzen besteht bei manchen Verlagen die Möglichkeit hauseigene Lizenzen zu vergeben, die sich am deutschen Urheberrecht orientieren.

5.5. Finanzierung

Die vorangegangenen Ausführungen zeigen ein sehr breites Dienstleistungsspektrum. Wie werden diese Services finanziert? Eingangs ist festzustellen, dass der Großteil der befragten Verlage an eine Institution nicht nur angebunden ist, wie es der Befragung zugrunde gelegt wurde, sondern keine eigenständige Rechtsform besitzt. Nur wenige Verlage sind als GmbH organisiert, ebenso wenige als Betrieb gewerblicher Art.

Bei der Finanzierung der Open-Access-Publikationen zeigt sich, dass zumeist eine Mischung unterschiedlicher Komponenten zum Einsatz kommt. Neben die Finanzierung durch Autor/inn/en treten dabei potentielle Verkaufserlöse der zusätzlich angebotenen Printexemplare oder eine Unterstützung durch die übergeordnete Institution vor allem im Hinblick auf Personalkosten und Infrastruktur; diese Kosten können vollständig oder teilweise übernommen werden. Drittmittel, Sponsoring und auch Verkaufserlöse aus Merchandise-Artikeln sind weitere Möglichkeiten zur Finanzierung.

5.6. Perspektiven

Die Verlage befinden sich in sehr unterschiedlichen Entwicklungsstadien, manche stehen noch ganz am Anfang, einige bestehen schon seit mehreren Jahren. Daher ist im Zusammenhang mit der Frage, wie die Verlage Open-Access-Monografien publizieren, interessant, welche Perspektiven sie dabei erwarten. Im Ergebnis lassen diese sich in vier Themenkomplexe gliedern: Dienstleistungen und Formate, Qualität und Reputation, Sichtbarkeit und Open Access, Wachstum und Entwicklung.

Was erwarten die Verlage dabei? Zum einen wurden Themen wie die Ausweitung des bestehenden Dienstleistungsangebots, der Ausbau der Online-Publikationen und die Einführung neuer Publikationsformate, wie EPUB, dynamische Publikationen oder auch die Möglichkeit medienneutraler Publikationsworkflows, genannt.

Daneben ist es für einzelne Verlage wichtig, ihr Renommee weiter auszubauen.

Publikationen sichtbar zu machen, ist ein wesentliches Anliegen des Open-Access-Publizierens. Daher sind die Universitätsverlage bemüht, die Sichtbarkeit ihrer Publikationen zu erhöhen. Maßnahmen hierzu sind beispielsweise die Registrierung bei einschlägigen Plattformen wie dem DOAB oder OAPEN. Ganz grundsätzlich sehen einige Befragte in diesem Bereich aber ebenso Ausbaubedarf im Hinblick auf ihre Open-Access-Publikationen.

Einen generellen Wachstumsbedarf lassen einige Verlage im Rahmen der Befragung erkennen, um das Fortbestehen des Verlages sichern zu können, aber auch um der bereits bestehenden Nachfrage besser gerecht werden zu können.

6. Diskussion

Die Universitätsverlage des deutschsprachigen Raums zeigen sich im Rahmen dieser Befragung sehr heterogen. Dies betrifft nicht nur den Zeitraum des Bestehens und die Anzahl der Publikationen. Insbesondere geht es im Hinblick auf Gold-Open-Access-Monografien um Fragen der Umsetzung von Open Access, des Wachstums oder auch der Weiterentwicklung. Diese Untersuchung kann aufgrund ihrer qualitativen Anlage nur einen Einblick und keinen absoluten Überblick bieten. Dennoch lassen sich aus den Ergebnissen einige Rückschlüsse gewinnen, wie sich die Situation der Universitätsverlage in ausgewählten Bereichen gestaltet.

6.1. Umgang mit Open Access

Unter den Begriff „Open Access“ wird oft Unterschiedliches verstanden, gerade im Hinblick auf die übertragenen und vorbehaltenen Rechte. Daher muss man sich bei der Themenstellung auch fragen, wie mit Gold Open Access für Monografien umgegangen wird. Betrachtet man konkret die Gruppe der Verlage, die tatsächlich ein Modell für Gold-Open-Access-Monografien anbieten, fallen zwei Aspekte auf: Was ist eigentlich unter Open Access zu verstehen und wie können überhaupt Autor/inn/en für Open Access gewonnen werden? Als Antwort auf den ersten Teil kann man feststellen: Die Auffassung scheint unterschiedlich zu sein. Als Open-Access-konform im Sinne der Berliner Erklärung würde man erwarten, dass die Verlage für Open-Access-Publikationen in der Regel eine Lizenz wie CC BY oder CC BY-SA vergeben, die dieses Kriterium erfüllen.⁵⁰ Es finden sich jedoch zahlreiche weitere Varianten in Form von CC-Lizenzen bzw. hauseigenen Lizenzen, die in der Regel kaum den Kriterien der Berliner Erklärung entsprechen. Diese Bandbreite an verschiedenen Lizenzen spiegelt sich auch in der Auswertung des DOAB bei E. Ferwerda, F. Pinter et al. wieder und beschränkt sich nicht nur auf den deutschsprachigen Raum.⁵¹

Gleichzeitig sind die Verlage an einigen Stellen mit der Problematik der Umsetzung von Open Access bei den Autor/inn/en konfrontiert. Dies zeigt sich speziell für Verlage, die sich in einem Umstellungsprozess zu Open Access befinden und nicht bereits als reine Open-Access-Verlage gegründet wurden. Vor allem wird dies am Anteil der Open-Access-Publikationen an den Gesamtpublikationen eines Verlages deutlich, der zwischen 0 und 100 % schwankt. Dabei kann der grüne Weg ebenfalls eine Lösung darstellen, um Autor/inn/en und Herausgeber/innen an das Konzept von Open Access heranzuführen.

50 „Lizenzen,“ Informationsplattform Open Access, zuletzt geprüft am 21.01.2018, <http://open-access.net/informationen-zu-open-access/rechtsfragen/lizenzen/>.

51 Ferwerda, Pinter und Stern, 39.

6.2. Autoren & Wachstum

Die Autor/inn/en als Inhaltslieferanten sind für die Universitätsverlage essentiell. Diese für eine Publikation im eigenen Verlag zu gewinnen, kann von Verlag zu Verlag unterschiedlich herausfordernd sein, ebenso sind die Wege hierzu verschieden. Viele Verlage setzen bei ihrer Vorgehensweise darauf, dass Autor/inn/en auf den Verlag aufmerksam werden und dem Verlag ihr Manuskript anbieten. Dieses eher passive Verhalten könnte für Verlage, die darauf abzielen ihre Publikationszahlen zu erhöhen, eher hinderlich sein. Andererseits weisen einige Verlage darauf hin, dass aufgrund knapper personeller Ressourcen eine Steigerung der Publikationszahlen kaum möglich ist. Hier zeigt sich, was ein weiterer Einflussfaktor für das Wachstum der Verlage sein kann. Mit Blick auf die unterschiedlichen Alter und Größen der Universitätsverlage wird aber deutlich, dass das Wachstumsziel nicht absolut für die Zweckerfüllung der Verlage ist. So bestehen einige Verlage über viele Jahre hinweg mit relativ kleinen Publikationszahlen.

6.3. Das Dienstleistungsangebot

Universitätsverlage bieten ein breites Spektrum an Verlagsdienstleistungen an, zu dieser Feststellung kam bereits A. Halle 2006.⁵² Dieses erweitert sich noch. Relevante Themen sind neben Lektorat, Druck und Vertrieb u.a. die Betreuung und Beratung von Autoren. Dies zeigt nicht nur das Ergebnis der Befragung, sondern z. B. auch ein Beitrag von R. Tobias zu KIT Scientific Publishing.⁵³ Darüber hinaus besteht ein Bedarf zur Weiterentwicklung der Services. Dies gilt speziell im Hinblick auf innovative Publikationsformate, die über die Veröffentlichung eines PDFs hinausgehen. Dies spiegelt zu einem gewissen Grad auch die sich verändernden Publikationsgewohnheiten.

6.4. Begutachtungsverfahren wissenschaftlicher Monografien

Zentral für wissenschaftliche Publikationen ist ihre wissenschaftliche Qualität. Die Absicherung der wissenschaftlichen Qualität der Publikationen erfolgt, wie es auch A. Halle beschreibt,⁵⁴ nach wie vor häufig durch Wissenschaftler/innen der eigenen Institution, beispielsweise bei der Begutachtung von Qualifikationsarbeiten. Den vielfach angestrebten Standard, ein Peer-Review-Verfahren durchzuführen, wie es in den Aufnahmekriterien des DOAB gefordert wird,⁵⁵ erfüllen nur wenige Verlage. Dass dieses Ziel für Monografien unter Umständen schwierig zu erreichen ist, da die fachlichen Gepflogenheiten unterschiedlich sind, räumt OAPEN ein.⁵⁶ Die Studie von K. Rosenbaum zeigt dasselbe,⁵⁷ gleichzeitig auch, dass der Aufwand für die Durchführung eines Peer-Review-Verfahrens sehr hoch ist und dabei auch zu zeitlichen Verzögerungen führen kann.

52 Axel Halle, „Universitätsverlage. Stand und Entwicklungsperspektiven,“ *Bibliotheksdienst* 40, Nr. 7 (2006): 809–817, <https://doi.org/10.1515/bd.2006.40.7.809>.

53 Vgl. Tobias, „Universitätsverlage sind ein wunderbares neues Geschäftsfeld für Bibliotheken“.

54 Axel Halle, „Wissenschaftliche Publikationskultur und Hochschulverlage,“ *Zeitschrift für Bibliothekswesen und Bibliographie* 50, Nr. 5 (2003): 243–250, zuletzt geprüft am 26.04.2018, http://zfbh.thulb.uni-jena.de/servlets/MCRFileNodeServlet/jportal_derivate_00001626/j03-h5-auf-1.pdf.

55 „Information for Publishers,“ Directory of Open Access Books, zuletzt geprüft am 21.01.2018, <http://www.doabooks.org/doab?func=about&uiLanguage=en>

56 „Description of the Peer Review process per partner,“ OAPEN, zuletzt geprüft am 21.01.2018, <http://oapen.org/content/peer-review-process-introduction>.

57 Rosenbaum, „Von Fach zu Fach verschieden.“

7. Fazit

Open Access trägt bei Monografien dazu bei, Erkenntnisse weiter zu verbreiten als dies in traditionellen Formaten möglich ist. Universitätsverlage haben dadurch eine Chance, ihrem Auftrag, Erkenntnisse zu verbreiten, in besonderer Weise gerecht zu werden. Für die Universitätsverlage des deutschsprachigen Raums zeigt sich, dass sich ihre Antworten auf die Fragen der Autorengewinnung, der Verbreitung, der Dienstleistungen, der Finanzierung des Angebots, der Qualitätssicherung und schließlich auch der konkreten Umsetzung von Open Access in vielerlei Hinsicht unterscheiden. Deutlich wird aber, dass einige Themen besonders wichtig sind, etwa die Frage, wie Open Access definiert wird; und darüber hinaus auch Themen des Publizierens an sich, z.B. wie Autor/inn/en für dieses Publikationsformat gewonnen werden können und wie Qualitätssicherung mit Unterstützung durch die Wissenschaftler/innen erfolgen kann und soll.⁵⁸

⁵⁸ Die Studie von Ferwerda, Pinter und Stern zieht ebenso den Schluss, dass Open-Access-Publizieren von Monografien auch eine Beschäftigung mit Fragen des Publizierens erfordert, die über Open Access hinaus gehen.

Anhang: Fragebogen

Open-Access-Monografien in Universitätsverlagen

Willkommen zur Umfrage „Open-Access-Monografien in Universitätsverlagen“

Schwerpunkt der Umfrage sind Autoren, Kunden, Geschäftsmodell und Qualitätssicherung für Open-Access-Monografien in Universitätsverlagen. Ziel der Masterarbeit ist eine Beschreibung der Universitätsverlage im deutschsprachigen Raum anhand dieser Bereiche und möglicher Entwicklungsperspektiven. Um einen breiten Einblick in das Feld mit einer einheitlichen Datenbasis zu bekommen, ist es wichtig, dass sich möglichst viele Universitätsverlage an der Befragung beteiligen.

Die Umfrage beinhaltet 15 Fragen. Die Antworten erfolgen, bedingt durch die qualitative Herangehensweise, vorwiegend in Form von Freitextfeldern. Eine Antwort in Form von Stichpunkten ist möglich. Sie können Eingaben speichern und zu einem späteren Zeitpunkt mit der Beantwortung fortzufahren.

Ende der Umfrage: 8. März 2017

Bei Fragen zu dieser Umfrage können Sie mich gerne kontaktieren: ...

Vielen Dank für Ihre Unterstützung!

Thema: Autoren & Kunden

Autoren und Leser sind wichtige Zielgruppen für wissenschaftliche Verlage. Bitte beantworten Sie im Folgenden einige Fragen zu Ihren Autoren und Lesern.

Bitte definieren Sie die Zielgruppen Ihres Verlags.

Wie gewinnen Sie Autoren für eine Publikation in Ihrem Verlag?

Welche Dienstleistungen bzw. Services bietet Ihr Verlag für Autoren an?

Für den Autor ist es wichtig, dass seine Arbeit gefunden und gelesen wird. Bitte beschreiben Sie, wie die Bücher bis hin zum Leser verbreitet und beworben werden.

Thema: Qualitätssicherung

Qualität ist ein Kriterium zur Bewertung von Dienstleistungen.

Wie stellt Ihr Verlag die inhaltliche Qualität der Manuskripte fest?

Neben der inhaltlichen Qualität ist auch die technische und formale Qualität der elektronischen bzw. der gedruckten Version wichtig. Wie sichert Ihr Verlag die formale und technische Qualität eines Buchs?

Wie sichern Sie eine gleichbleibende Qualität der Dienstleistung, z.B. im Sinne eines Qualitätsmanagements?

Thema: Geschäftsmodell

Welche Organisationsform hat Ihr Verlag?

Wie deckt Ihr Verlag die ihm entstehenden Kosten, beispielsweise Druckkosten, Personal, Infrastruktur? Bitte berücksichtigen Sie hier ggf. auch Kosten und Infrastruktur (z. B. Räume), die von Ihrer Einrichtung bereitgestellt werden.

Wie handhabt Ihr Verlag das Thema Open Access? Welche Lizenzen bieten Sie an?

Welche Perspektiven sehen Sie für die Weiterentwicklung Ihres Universitätsverlags?

Angaben zum Verlag

Um Ihre Antworten in den richtigen Kontext einordnen zu können, bitte ich Sie zum Abschluss der Umfrage um einige allgemeine Angaben.

Name des Verlags:

Gründungsjahr des Verlags:

Wie viele Bücher hat Ihr Verlag in den vergangenen Jahren veröffentlicht?

Anzahl der Bücher insgesamt	davon Open Access
2016	
2015	
2014	
2013	
2012	

Möchten Sie noch etwas ergänzen?

Einwilligungserklärung zur Verwendung der Daten

Mit dem Absenden der Umfrage erklären Sie sich einverstanden, dass die Daten für meine Masterarbeit genutzt werden und eine Einsichtnahme durch die Gutachter erfolgen kann. Die im Rahmen dieser Befragung erhobenen Daten werden vorübergehend auf einem Server der Humboldt-Universität gespeichert und nach Abschluss der Befragung gelöscht.

Bitte wählen Sie im Folgenden aus, welchen darüberhinausgehenden Nachnutzungsmöglichkeiten der Daten Sie zustimmen.

Bei Rückfragen oder für eine Änderung Ihrer Einwilligung erreichen Sie mich per E-Mail:

Im Falle einer (Open-Access-)Veröffentlichung der Masterarbeit, stimme ich der Verwendung der Daten zu.

Ja / In anonymisierter Form

Einer Nachnutzung der Daten für wissenschaftliche Zwecke und gegebenenfalls durch Teilnehmer der Befragung stimme ich zu. *

Ja / Nur in anonymisierter Form / Nein

Einer Veröffentlichung der Daten als Open Data stimme ich zu. *

Ja / Nur in anonymisierter Form / Nein

Vielen Dank für Ihre Antwort!

Literaturverzeichnis

- Bargheer, Margo und Jutta Pabst. „Being small is not a fault’: Making sense of the newer generation of German-language university presses.“ *Learned Publishing* 29 (2016): 335–341. <https://doi.org/10.1002/leap.1053>.
- Eve, Martin P. „Open Access publishing and scholarly communications in non-scientific disciplines.“ *Online Information Review* 39, Nr. 5 (2015): 717–732. <https://doi.org/10.1108/OIR-04-2015-0103>.
- Ferwerda, Eelco. „Open access monograph business models.“ *Insights: the UKSG journal* 27, s1 (2014): 35–38. <https://doi.org/10.1629/2048-7754.46>.
- Ferwerda, Eelco, Frances Pinter und Niels Stern. *A Landscape Study On Open Access And Monographs: Policies, Funding And Publishing In Eight European Countries*. 2017. <https://doi.org/10.5281/zenodo.815932>.
- Fournier, Johannes und Roland Weihberg. „Das Förderprogramm »Open Access Publizieren« der Deutschen Forschungsgemeinschaft. Zum Aufbau von Publikationsfonds an wissenschaftlichen Hochschulen in Deutschland.“ *Zeitschrift für Bibliothekswesen und Bibliographie* 60, Nr. 5 (2013): 236–243. <https://doi.org/10.3196/186429501360528>.
- Fund, Sven. „De Gruyter Open Library – Das Open Access-Modell eines Universalwissenschaftsverlags für Bücher und Zeitschriften.“ *BIBLIOTHEK Forschung und Praxis* 35, Nr. 2 (2011): 157–160. <https://doi.org/10.1515/bfup.2011.021>.
- Fund, Sven. „Will Open Access Change the Game?“ *BIBLIOTHEK Forschung und Praxis* 39, Nr. 2 (2015): 206–209. <https://doi.org/10.1515/bfp-2015-0025>.
- Georg-August-Universität Göttingen. *Richtlinie für den Universitätsverlag Göttingen* (09.01.2015). Zuletzt geprüft am 21.01.2018. <https://www.univerlag.uni-goettingen.de/dokumente/texts/Nutzungsordnung-2015-01-09.pdf>.
- Givler, Peter. „University Press Publishing in the United States.“ In *Scholarly Publishing: Books, Journals, Publishers and Libraries in the Twentieth Century*, herausgegeben von Richard E. Abel und Lyman W. Newman. Wiley, 2002. <http://www.aaupnet.org/about-aaup/about-university-presses/history-of-university-presses>.
- Halle, Axel. „Wissenschaftliche Publikationskultur und Hochschulverlage.“ *Zeitschrift für Bibliothekswesen und Bibliographie* 50, Nr. 5 (2003): 243–250. Zuletzt geprüft am 26.04.2018. http://zfbib.thulb.uni-jena.de/servlets/MCRFileNodeServlet/jportal_derivate_00001626/j03-h5-auf-1.pdf.

- Halle, Axel. „Universitätsverlage. Eine vergleichende Perspektive.“ *Zeitschrift für Bibliothekswesen und Bibliographie* 51, 5/6 (2004): 277–283. Zuletzt geprüft am 26.04.2018. http://zfbib.thulb.uni-jena.de/servlets/MCRFileNodeServlet/jportal_derivate_00001709/j04-h5-6-auf-2.pdf.
- Halle, Axel. „Universitätsverlage. Stand und Entwicklungsperspektiven.“ *Bibliotheksdienst* 40, Nr. 7 (2006): 809–817. <https://doi.org/10.1515/bd.2006.40.7.809>.
- Häußermann, Sabine. „Aspekte der Gründung eines Universitätsverlages am Beispiel Heidelberg.“ *Berliner Handreichungen zur Bibliotheks- und Informationswissenschaft* 255. Berlin: Institut für Bibliotheks- und Informationswissenschaft der Humboldt-Universität zu Berlin, 2009. <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:kobv:11-10099919>.
- Huber, Verena. „Zur Typologisierung der aktuellen deutschen Verlagslandschaft.“ *Alles Buch: Studien der Erlanger Buchwissenschaft* 46. Friedrich-Alexander-Universität Erlangen-Nürnberg, 2012. <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:29-opus4-58742>.
- Keller, Alice. „Publikationskompetenz als neues Tätigkeitsfeld von Bibliotheken.“ *Bibliotheksdienst* 50, Nr. 7 (2016): 661–671. <https://doi.org/10.1515/bd-2016-0080>.
- Kuckartz, Udo. *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. 3., überarbeitete Auflage. *Grundlagentexte Methoden*. Beltz, 2016.
- Lucius, Wulf D. von. *Verlagswirtschaft: ökonomische, rechtliche und organisatorische Grundlagen*. 3., neubearbeitete und erweiterte Aufl. Konstanz: UVK-Verlagsgesellschaft, 2014.
- Mudditt, Alison. „The Past, Present, and Future of American University Presses: A View from the Left Coast.“ *Learned Publishing* 29 (2016): 330–334. <https://doi.org/10.1002/leap.1047>.
- Pflug, Günther. „Universitätsverlag.“ In *Lexikon des gesamten Buchwesens: Band VII*, herausgegeben von Severin Corsten et al. 2., völlig neu bearbeitete und erweiterte Aufl., 596. Stuttgart: Hiersemann, 2007.
- Pinter, Frances. „Open Access for Scholarly Books?“ *Publishing Research Quarterly* 28, Nr. 3 (2012): 183–191. <https://doi.org/10.1007/s12109-012-9285-0>.
- Rosenbaum, Konstanze. „Von Fach zu Fach verschieden. Diversität im wissenschaftlichen Publikationssystem.“ In *Wissenschaftliches Publizieren: Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung*, herausgegeben von Peter Weingart und Niels Taubert, 41–74. <https://doi.org/10.1515/9783110448115-002>.

- Schallehn, Volker und Ralf Schimmer. „Open Access.“ In *Praxishandbuch Bibliotheksmanagement*, herausgegeben von Rolf Griebel et al., 311–38. Reference. Berlin: de Gruyter Saur, 2016. <https://doi.org/10.1515/9783110303261.311>.
- Schönstedt, Eduard und Thomas Breyer-Mayländer. *Der Buchverlag: Geschichte, Aufbau, Wirtschaftsprinzipien, Kalkulation und Marketing*. Stuttgart: Metzler, 2010.
- Taubert, Niels. „Open Access und digitale Publikation aus der Perspektive von Wissenschaftsverlagen.“ In *Wissenschaftliches Publizieren: Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung*, herausgegeben von Peter Weingart und Niels Taubert, 75–102. <https://doi.org/10.1515/9783110448115-003>.
- Taubert, Niels. „Formale wissenschaftliche Kommunikation.“ In *Forschungsfeld Wissenschaftskommunikation*, herausgegeben von Heinz Bonfadelli et al., 125–39. Wiesbaden: Springer, 2017. https://doi.org/10.1007/978-3-658-12898-2_7.
- Thatcher, Sanford G. „From the University Presses: What University Presses Think about Open Access.“ *Against the Grain* 22, Nr. 3 (2010): 55–58. <https://doi.org/10.7771/2380-176X.5570>.
- Tobias, Regine. „Universitätsverlage sind ein wunderbares neues Geschäftsfeld für Bibliotheken.“ *BIT online* 18, Nr. 2 (2015): 170–176. Zuletzt geprüft am 21.01.2018. <http://b-i-t-online.de/heft/2015-02-interview-tobias.pdf>.
- Weingart, Peter and Niels Taubert, Hrsg. *Wissenschaftliches Publizieren: Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung*. Berlin: de Gruyter, 2016.
- Wissenschaftsrat, *Empfehlungen zur digitalen Informationsversorgung durch Hochschulbibliotheken*. Greifswald, 2001. Zuletzt geprüft am 21.01.2018. <http://www.wissenschaftsrat.de/download/archiv/4935-01.pdf>.
- Woll, Christian. *Bibliotheken als Dienstleister im Publikationsprozess: Herausforderungen und Chancen alternativer Formen des wissenschaftlichen Publizierens*. Saarbrücken: VDM Verlag Müller, 2006.

Zeitretrieval in der verbalen Inhaltserschließung

Esther Scheven, Deutsche Nationalbibliothek, Frankfurt am Main

Michael Wening, Öffentliche Bibliothek, Remscheid

Zusammenfassung:

Die Zeitdimension spielt im Retrieval oft eine wichtige Rolle. Dennoch ist eine inhaltliche Suche nach der Zeitfacette in Katalogen und anderen Rechtersystemen kaum sinnvoll umgesetzt, da die zugrundeliegenden Erschließungssysteme die Erfassung der Zeitfacette nur unzureichend abbilden. Der Artikel stellt kurz den Status quo dar und entwickelt eine Idee für ein verbessertes Retrieval der Zeitschlagwörter bei der inhaltlichen Erschließung.

Summary:

Time plays an important role in bibliographic retrieval. Despite of that, the results of subject-focused time searches in library catalogs and other retrieval systems are often inadequate, as the cataloging and indexing systems do not sufficiently represent the time aspect. The article explains the status quo and goes on to discuss an idea for an improved recording and retrieval of temporal subject headings.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S181-192>

Autorenidentifikation: Scheven, Esther: GND 11399849X; Wening, Michael: GND 1149224487

Schlagwörter: Inhaltserschließung; Zeitfacette; Retrieval

1. Zeitaspekte in Formal-, Provenienz- und Inhaltserschließung

Das Phänomen Zeit ist seit jeher ein großes Faszinosum, welches nicht nur in der Alltagserfahrung der Menschen eine wichtige Rolle spielt, sondern auch in Philosophie, Wissenschaft, Literatur und Kunst Gegenstand der Untersuchung und Reflexion ist.

Zeitaspekte werden in allen drei bibliothekarischen Erschließungsarten relevant: Zeitangaben in der Formalerschließung erfassen und beschreiben die formalen zeitlichen Merkmale, welche geeignet sind, um Ressourcen in einen temporalen Kontext ihrer Entstehung bzw. Erscheinung einzuordnen. Dazu gehört auch die Erfassung von Lebensdaten bei Personen oder Zeitangaben wie Gründungsjahr oder Zeitpunkt der Veranstaltung bei Körperschaften und Konferenzen. Zeitangaben in der Provenienzerschließung dienen dazu, Aussagen über persönliche sowie körperschaftliche Vorbesitzer einzelner Ressourcen zu treffen, Provenienzketten zu rekonstruieren und dadurch die jeweils eigenständige Geschichte eines Buches als physisches Objekt zu beleuchten.

Die zeitlichen Inhalte einer Ressource, die im Zentrum der folgenden Überlegungen stehen, werden mittels klassifikatorischer und verbaler Erschließung in Form von Notationen beziehungsweise Deskriptoren erfasst; sie sollen Aufschluss darüber geben, wie das Thema eines Dokumentes chronologisch einzuordnen ist.

Viele Klassifikationen verfügen über Systemstellen, die einer zeitlichen Gliederung folgen (Abb. 1, DDC-Notation 943.073 „1815–1847“; Unterteilung in historische Epochen), oder sehen – wo eine Einordnung von Ressourcen in diese Klassen nicht möglich ist – die Verwendung von Zeitschlüsseln (Abb. 2, DDC Hilfstafel 1, Notation-0903 „Neuzeit, 1500–“; Unterteilung in Jahrhunderte) im Rahmen synthetischer Notationen vor (beispielsweise Dewey-Dezimalklassifikation [DDC]¹).

The screenshot shows a hierarchical view of the DDC class 943.073 1815-1847. The root node is '943.073 1815-1847'. It branches into several sub-nodes: '900' (Geschichte & Geografie), '930-990' (Geschichte einzelner Kontinente, Länder, Ortschaften; außerirdischer Welten), '940-990' (Geschichte einzelner Kontinente, Länder, Ortschaften der Neuzeit; Geschichte außerirdischer Welten), '940' (Geschichte Europas), '943' (Deutschland und benachbarte mitteleuropäische Länder), '943.01-943.08' (Zeitabschnitte der Geschichte Deutschlands), and '943.07' (Zeit des Deutschen Bundes, 1815-1866). The selected node '943.073 1815-1847' is highlighted in a grey bar at the bottom of the list.

Abb. 1. WebDewey Deutsch, Beispiel für Systemstelle mit zeitlicher Gliederung²

The screenshot shows a hierarchical view of the DDC class T1--0903 *Neuzeit, 1500-. The root node is 'T1--0903 *Neuzeit, 1500-'. It branches into several sub-nodes: 'T1--0' (Hilfstafel 1. Standardschlüssel), 'T1--09' (Geschichte, geografische Behandlung Biografien), and 'T1--0901-T1--0905' (Zeitabschnitte). The selected node 'T1--0903 *Neuzeit, 1500-' is highlighted in a grey bar. Below this bar, a list of sub-nodes is shown: 'T1--09031' (*16. Jahrhundert, 1500–1599), 'T1--09032' (*17. Jahrhundert, 1600–1699), 'T1--09033' (*18. Jahrhundert, 1700–1799), and 'T1--09034' (*19. Jahrhundert, 1800–1899).

Abb. 2. WebDewey Deutsch, Zeitschlüssel aus der DDC Hilfstafel 1³

1 Dewey-Dezimalklassifikation (Deutsch), zuletzt geprüft am 28.03.2018, <https://deweyde.pansoft.de/webdewey/login/login.html>.

2 Abbildung: OCLC, CC BY-NC-ND 3.0.

3 Abbildung: OCLC, CC BY-NC-ND 3.0.

Normalerweise werden Notationen aus den Hilfstafeln gemäß den Anweisungen zu einer synthetischen Notation zusammengefügt. Da die meisten Nutzer verbal suchen, hat die Deutsche Nationalbibliothek (DNB) mit Einführung der DDC zusätzlich zur synthetischen Notation die Notationsbestandteile einzeln abgelegt. Über ein spezielles Tool WebDewey Search⁴ lassen sich DDC-Notationen verbal suchen. Aufgrund der genannten Einzelablage der DNB lässt sich der Zeitschlüssel von mit DDC erschlossenen Dokumenten separat suchen.⁵

Bei der verbalen Erschließung nach den „Regeln für die Schlagwortkatalogisierung (RSWK)“⁶ dienen in geringem Umfang Epochenbezeichnungen sowie als maßgebliches Instrumentarium die Zeitschlagwörter der zeitlichen Indexierung von Dokumenten. Bei Epochenbezeichnungen wie Barock oder Renaissance ist der Zeitaspekt immanent und wird im Normdatensatz erfasst.⁷ Beispielsweise ist für die Epochenbezeichnung „Barock“ die Zeitspanne von 1570 bis 1750 im Normdatensatz angegeben. Aus den Verwendungshinweisen geht überdies hervor, dass bei Werken aus dem Bereich der Kunst die gültige Zeitspanne auf den Zeitraum von 1600 bis 1750 verkürzt ist.⁸ Epochenbezeichnungen werden bei der Schlagwortvergabe wie Sachschlagwörter behandelt.

Die nach RSWK zugelassenen acht Zeitschlagwörter⁹ haben einen eigenen Indikator und stehen bei der Bildung von Schlagwortfolgen in der Regel am Ende. Üblicherweise sind sie um Jahresangaben ergänzt. Nur bei Gesamtdarstellungen fehlen diese. Problematisch ist bisher die mangelnde Verwertbarkeit der zeitlichen Angaben aus den Normdatensätzen und Zeitschlagwörtern für das Retrieval in elektronischen Katalogen.

Beispiele für Schlagwortfolgen mit Epochenbezeichnung und Zeitschlagwort:

Titel: Grundriss Philosophie des Humanismus und der Renaissance (1350–1600) / Thomas Lein-kauf. – 2017

Schlagwortfolgen: Renaissance ; Humanismus ; Philosophie
Philosophie ; z Geschichte 1350–1600¹⁰

Titel: Untersuchungen zur römischen Keramik und weiteren Funden aus Syene/Assuan (1.–7. Jahrhundert AD) : Grabungen 2001–2004 / von Stefanie Martin-Kilcher und Jacqueline Winingler. – 2017

Schlagwortfolge: Assuan ; Funde ; z Geschichte 1–700¹¹

4 WebDeweySearch, zuletzt geprüft am 28.03.2018, <https://deweysearchde.pansoft.de/webdeweysearch>.

5 WebDeweySearch, zuletzt geprüft am 28.03.2018, <https://deweysearchde.pansoft.de/webdeweysearch>. Im Portal der DNB: „Browsen (DDC)“, Hilfstafel und Zeitabschnitt aussuchen; Treffermenge im Katalog auf die Sachgruppe einschränken.

6 Deutsche Nationalbibliothek, *Regeln für die Schlagwortkatalogisierung*: RSWK, 4. vollständig überarbeitete Auflage 2017, Stand: März 2017, [urn:nbn:de:101-2017011305](https://nbn-resolving.org/urn:nbn:de:101-2017011305).

7 Im Feld 548, siehe „GND-Erfassungsleitfaden, Feld 548,“ zuletzt geprüft am 28.03.2018, <https://wiki.dnb.de/download/attachments/50759357/548.pdf>.

8 Datensatz der Deutschen Nationalbibliothek, zuletzt geprüft am 28.03.2018, <http://d-nb.info/gnd/4004541-9>.

9 Die acht Zeitschlagwörter nach RSWK: §402 lauten: Geschichte, Prognose, Geistesgeschichte, Ideengeschichte, Kirchengeschichte, Sozialgeschichte, Vor- und Frühgeschichte, Weltgeschichte.

10 Datensatz der Deutschen Nationalbibliothek, zuletzt geprüft am 28.03.2018, <http://d-nb.info/1130678040>.

11 Datensatz der Deutschen Nationalbibliothek, zuletzt geprüft am 28.03.2018, <http://d-nb.info/1149778997>.

2. Retrieval der Zeitdimension – Überblick über die Regelwerkssituation und die verwendeten Normdateien (2003–2017)

2003 hatte sich Jutta Frommeyer in ihrer Dissertation „Zeitbegriffe und Zeitcodierungen in allgemeinbibliographischen Datenbanken“ dem Thema „Zeitretrieval“ gewidmet.¹² Sie untersuchte die „Regeln für den Schlagwortkatalog (RSWK)“,¹³ die Library of Congress Subject Headings (LCSH) als Regelwerk sowie das französische Pendant und die nach diesen Regeln erstellten Normdateien (SWD, LCSH, RAMEAU) und stellte ein theoretisches Modell für ein Zeitretrieval vor mit den Elementen Zeitcode, Chronikcode und Chroniknormdatei. Der Zeitcode ist ein numerisches Feld in der bibliografischen Beschreibung; Zeitangaben werden als Zahlen und nicht als Zeichen erfasst, der Chronikcode ist analog zum Zeitcode ein numerisches Feld innerhalb eines Normdatensatzes für die Erfassung des zeitlichen Bezugs. Bei der Chroniknormdatei handelt es sich um eine eigene Normdatei für alle Normdatensätze mit einem zeitlichen Bezug ausgehend von der zeitlichen Komponente, d.h. es werden einzelne Normdatensätze für die zeitliche Komponente gebildet, die die Vorzugsbenennung darstellt. Die Erfassung erfolgt in der gleichen Struktur wie des Zeitcodes.¹⁴

Beispiele für Normdatensätze einer Chroniknormdatei nach dem Vorschlag von Jutta Frommeyer (Grundlage sind Datensätze aus der GND; die Indikatoren entsprechen den Absprachen in den RSWK¹⁵):

- 1939–1945 g Tempelburg (und viele weitere in der GND)
 k Andreas-Hofer-Bund (und viele weitere in der GND)
 p Voigtel, Heinrich (Wirkungsjahre) (und weitere in der GND)
 s Arbeitserziehungslager
 s Panzer III
 s Weltkrieg (1939-1945) (und einige weitere in der GND)
 t Wehrmachtsbericht (und viele weitere in der GND)
- 2004-01-01/2012-09-30 k Agentur für Arbeit Koblenz
 k Agentur für Arbeit Mayen

Aufgrund der gleichen Struktur der Zeitangaben in der bibliografischen Beschreibung, in den Zeiterfassungsfeldern von Normdatensätzen und den speziellen Normdatensätzen der Chroniknormdatei können umfassende Suchen sowohl mit den Zeitangaben als auch mit verbalen Entsprechungen durchgeführt werden.

12 Frommeyer, Jutta, *Zeitbegriffe und Zeitcodierungen in allgemeinbibliographischen Datenbanken: Eine Analyse von RSWK, LCSH und RAMEAU sowie Entwicklung eines theoretischen Modells für ein Zeitretrieval* (Berlin: Logos-Verl., 2003).

13 Bis zur dritten Auflage hieß der Titel „Regeln für den Schlagwortkatalog“; mit der vierten Auflage wurde er in „Regeln für die Schlagwortkatalogisierung“ geändert.

14 Esther Scheven, „Rezension zum Buch von Jutta Frommeyer: Zeitbegriffe und Zeitcodierungen in allgemeinbibliographischen Datenbanken,“ *Zeitschrift für Bibliothekswesen und Bibliographie* 50 (2003): 295–297, zuletzt geprüft am 28.03.2018, http://zfbib.thulb.uni-jena.de/servlets/MCRFileNodeServlet/jportal_derivate_00001835/j03-h5-rez-1.pdf.

15 Indikatoren: g = Geografika; k = Körperschaften; p = Personen; s = Sachbegriff; t = Werknormsatz.

Zwischenzeitlich haben sich bibliothekarische Modelle und Regelwerke sowie die im D-A-CH-Raum verwendete Gemeinsame Normdatei (GND) weiterentwickelt. Allgemein bekannt ist die FRBR-Familie,¹⁶ die mit dem IFLA Library Reference Model (IFLA LRM)¹⁷ wohl zunächst ihren Abschluss gefunden hat. FRBR und IFLA LRM verwirklichen ein Entity-Relationship-Model, das heißt, verschiedene Entitäten werden zueinander in Beziehung gesetzt und die Art der Beziehung gekennzeichnet. Das neue Regelwerk Resource Description and Access (RDA)¹⁸ gründet sich auf FRBR beziehungsweise zukünftig auf IFLA LRM. Bei der Entwicklung der Gemeinsamen Normdatei (GND) wurde auch auf ein Entity-Relationship-Model gesetzt. RSWK ist nicht mehr länger das Regelwerk für *alle* Entitäten der Sacherschließung, die nun in der GND vorgehalten werden. Die Regeln zur Erstellung von Normdatensätzen der GND für mögliche geistige Schöpfer (Personen, Körperschaften) und Werke, die auch Thema sein können, folgen seit dem Umstieg des D-A-CH-Raums auf RDA dem neuen Regelwerk.

Beim Ursprungsmodell FRBR ist die Zeitfacette in der Gruppe 3 in Form der Entität „Event“ dargestellt, jedoch ausschließlich für inhaltliche Aspekte der Erschließung vorgesehen. Die in den 2000er-Jahren erarbeiteten Modifizierungen von FRBR versuchten in der Folge, die Zeitentität unabhängig von der Erschließungsart zu definieren und den Informationseinrichtungen somit einen größeren Spielraum in der Gestaltung ihrer spezifischen Regelwerke und Recherchesysteme zu belassen. Im LRM ist die Entität „Ereignis“ in „Zeitspanne“ („Time-span“, LRM-E11) umbenannt worden und rangiert zusammen mit den meisten anderen Entitäten auf der zweiten Hierarchieebene des Referenzmodells. Damit ist die Entität „Zeitspanne“ eine Instanz der Entität „Res“ und kann für jedwede Art von Zeitangabe verwendet werden.

Nicht zuletzt aufgrund dieser tiefgreifenden Veränderungen und weil der Standardisierungsausschuss eine Weiterentwicklung der RSWK intendiert, erschien es sinnvoll, im Rahmen einer Masterarbeit die Zeitdimension in der bibliothekarischen inhaltlichen Erschließung erneut zu untersuchen und Empfehlungen für ein verbessertes Retrieval auszuarbeiten. Die Masterarbeit von Michael Wening,¹⁹ die an der DNB erstellt wurde, analysiert die Vorgaben für die Erfassung der Zeitfacette in klassifikatorischen und verbalen Erschließungsregelwerken und bewertet sie bezogen auf das Retrieval.

Wie er herausgearbeitet hat, ist beispielsweise ein großer Nachteil von RDA, dass Zeitangaben kaum maschinenlesbar erfasst werden. In RDA 6.20.3.3, wo es um die Erfassung des Datums eines Abkommens geht, heißt es: „Erfassen Sie das Datum in der Form [Jahr] [Monat] [Tag]. Erfassen Sie

16 FRBR-Familie, siehe „IFLA Conceptional Models,“ zuletzt geprüft am 28.03.2018, <https://www.ifla.org/node/2016>. „FRBR = Functional Requirements for Bibliographic Records (1998),“ zuletzt geprüft am 28.03.2018, <https://www.ifla.org/publications/functional-requirements-for-bibliographic-records>. „FRAD = Functional Requirements for Authority Data (2009),“ zuletzt geprüft am 28.03.2018, <https://www.ifla.org/publications/functional-requirements-for-authority-data>. „FRSAD = Functional Requirements for Subject Authority Data (2011),“ zuletzt geprüft am 28.03.2018, <https://www.ifla.org/node/5849>.

17 „IFLA Library Reference Model (IFLA LRM) (2017),“ zuletzt geprüft am 28.03.2018, <https://www.ifla.org/publications/node/11412>.

18 Resource Description and Access, zuletzt geprüft am 28.03.2018, <http://access.rdatoolkit.org/index.php>.

19 Michael Wening, *Die Zeitfacette in bibliothekarischen Erschließungen: Eine Analyse unter besonderer Berücksichtigung von sachlicher Erschließung und Retrieval*, Masterarbeit im Rahmen des Masterstudiengangs „Bibliotheks- und Informationswissenschaft“ (Berlin: Humboldt-Universität, 2017), zuletzt geprüft am 28.03.2018, <http://d-nb.info/1149163542>.

den Monat in der Sprache und Schrift, die die Agentur bevorzugt, welche die Daten erzeugt.“ Das heißt, dass je nach Arbeitssprache ganz unterschiedliche Angaben für den Monat gemacht werden (Juni – June – Juin – Junio etc.). Um eine bessere internationale Standardisierung der Katalogdaten erreichen zu können, müssten Zeitangaben im besten Fall in einem rein numerischen Datumsformat erfasst werden. Zeitangaben sind nicht nur in der bibliothekarischen Welt wichtig, sondern in vielen Zusammenhängen. Die Anforderungen sind für alle Anwendergemeinschaften ähnlich. Deshalb ist es sinnvoll, solche Angaben über allgemeine Standards wie Normen des DIN oder der internationalen Normungsorganisation ISO zu regeln. In der internationalen Norm ISO 8601²⁰ werden Empfehlungen zur Erfassung von Zeitangaben gegeben. Die bibliothekarischen Standards sollten die allgemeinen Standards berücksichtigen und ggf. ganz auf sie verweisen, wenn sie für die Erfassung von Katalogdaten ausreichend sind.

Ebenfalls untersucht wurden die angebotenen Möglichkeiten für ein Zeitretrieval in großen Bibliothekskatalogen (Verbundkataloge, Katalog der DNB), wobei auch Vergleiche mit Spezialportalen²¹ angestellt wurden, weil diese für ihren speziellen Focus häufig andere oder elaborierte Zugänge anbieten. Wie sich zeigte, hat sich an den Möglichkeiten der inhaltlichen Suche nach der Zeitfacette seit 2003 in den Bibliothekskatalogen kaum etwas verbessert. Klassifikatorisch erschlossene Titel sind bezogen auf den Zeitaspekt i.d.R. schwer recherchierbar, da die Zeitfacette üblicherweise Teil der Gesamtnotation ist. Durch die Einzelablage der verschiedenen Bestandteile von synthetischen Notationen der DNB kann über WebDewey Search zwar die Zeitfacette gesucht werden; die Suche ist aber nicht intuitiv.²² Die Erfassung der Zeitdimension nach RSWK ist nicht normiert; es werden immer die konkreten Anfangs- und Endjahre erfasst. Die erfassten Zeitabschnitte sind für den Endnutzer nicht vorhersehbar und deshalb im Retrieval schlecht zu verwenden. Im Rahmen der verbalen Erschließung hat Michael Wening diskutiert, ob die Erfassung von normierten Zeitabschnitten oder Normdatensätze für Zeitschlagwörter eine Verbesserung bringen würden.

Der von Jutta Frommeyer geforderte Chronikcode ist in der GND durch die neue Datenstruktur verwirklicht: Im Feld 548 sollen bei allen Normdatensätzen mit Zeitbezug die Angaben strukturiert erfasst werden.

Beispiele, Katalogdarstellung in PICA3:

150 Renaissance

548 1420\$b1600\$4datb²³

20 „ISO 8601 Date and time format, Data elements and interchange formats – Information interchange – Representation of dates and times,“ zuletzt geprüft am 28.03.2018, <https://www.iso.org/standard/40874.html>. Neue Entwürfe von 2017: „Teil 1: Basic rules,“ zuletzt geprüft am 28.03.2018, <https://www.iso.org/obp/ui#iso:std:iso:8601:-1:dis:ed-1:v1:en>. „Teil 2: Extensions,“ zuletzt geprüft am 28.03.2018, <https://www.iso.org/obp/ui#iso:std:iso:8601:-2:dis:ed-1:v1:en>.

21 Eine Darstellung des Zeitretrievals der untersuchten Spezialportale sprengt den Rahmen von diesem Artikel. Deshalb sei auf die Ausführungen in der Masterarbeit verwiesen: Wening, *Die Zeitfacette*, 34–37.

22 Perspektivisch wird die DNB die Erfassung von synthetischen Notationen mit Einzelablage durch die Vergabe von Kurznotationen ersetzen.

23 Datensatz der Deutschen Nationalbibliothek, zuletzt geprüft am 28.03.2018, <http://d-nb.info/gnd/4049450-0>.

110 Deutsch-Finnische Gesellschaft

510 !98806328X!Deutsch-Finnische Gesellschaft in Süddeutschland [Tb1]\$4vorg
548 1956\$4datb²⁴

100 Schiller, Friedrich

548 1759\$b1805\$4datl
548 10.11.1759\$b09.05.1805\$4datx²⁵

Eine Chroniknormdatei, in der ausgehend von der Zeitangabe alle Normdatensätze mit der gleichen Zeitangabe zusammengefasst werden, gibt es zurzeit allerdings in der Weise nicht. Bei der Entwicklung der GND wurde davon Abstand genommen, Normdatensätze für Zeitangaben zu erfassen, da solche Datensätze ins Uferlose anwachsen und bei guten Retrievalmöglichkeiten auch nicht gebraucht würden.²⁶ Stattdessen empfiehlt Michael Wening, für das Zeitretrieval nicht nur die Angaben mit Zeitbezug in der Schlagwortfolge, sondern auch etwaige Zeitangaben in den Normdatensätzen der übrigen Schlagwörter in der Schlagwortfolge auszuwerten. Das würde bei einer zeitlichen Suche die relevante Treffermenge merklich erhöhen.

Eine *rasche* Verbesserung im Retrieval ließe sich bei einer verbesserten Auswertung der Zeitschlagwörter nach RSWK erreichen. Die Erfassung in einem neuen Zeitcode-Feld in der bibliografischen Beschreibung analog zum Vorschlag von Jutta Frommeyer ist wahrscheinlich nicht notwendig; eine bessere Strukturierung der derzeit erfassten Zeitangaben aber sinnvoll. Im Rahmen der Masterarbeit von Michael Wening wurde ein Vorschlag für ein verbessertes Retrieval der Zeitschlagwörter nach RSWK erarbeitet, der im Folgenden vorgestellt wird.

3. Recherche nach Zeitschlagwörtern der RSWK und mögliche Weiterentwicklung

Es lassen sich zwei grundsätzliche Arten von Zeitschlagwörtern unterscheiden: Zeitschlagwörter ohne und mit Jahreszahlerweiterung. Zeitschlagwörter ohne Jahreszahlerweiterung sind zu verwenden, wenn es sich um Gesamtdarstellungen handelt, d.h. wenn die gesamte Geschichte der thematisierten Sache dargestellt wird. Bei Inhalten mit zeitlicher Begrenzung (d.h. wenn es um einen bestimmten zeitlichen Ausschnitt geht) wird das Zeitschlagwort mit den entsprechenden Jahreszahlen vergeben.²⁷

Beispiele:

Zeitschlagwort ohne Jahreszahlerweiterung, da Gesamtdarstellung

Titel: Der Dreißigjährige Krieg / Georg Schmidt. – 2010²⁸

24 Datensatz der Deutschen Nationalbibliothek, zuletzt geprüft am 28.03.2018, <http://d-nb.info/gnd/1029712-1>.

25 Datensatz der Deutschen Nationalbibliothek, zuletzt geprüft am 28.03.2018, <http://d-nb.info/gnd/118607626>.

26 Der Gedanke von Jutta Frommeyer, über eine Chroniknormdatei sprachlich unterschiedliche Normdateien (beispielsweise LSCH, RAMEU, GND) miteinander zu verknüpfen, sollte weiterhin verfolgt werden.

27 Vergleiche Paragraph 401a, Nr. 2, a) + b), RSWK.

28 Datensatz der Deutschen Nationalbibliothek, zuletzt geprüft am 28.03.2018, <http://d-nb.info/1001926684>.

Schlagwörter: Dreißigjähriger Krieg ; Geschichte
Behandelt ist der Gesamtzeitraum von 1618–1648

Zeitschlagwort mit Jahreszahlerweiterung, da die Darstellung zeitlich begrenzt ist:
Titel: Sozialgeschichte Baden-Württembergs : 1800–1989 / Willi A. Boelcke. – 1989²⁹
Schlagwörter: Baden-Württemberg ; Sozialgeschichte 1800–1989

Im Folgenden sollen die Zeitschlagwörter mit erweiterter Jahreszahlangabe näher betrachtet werden.³⁰ Deren wesentlicher Nachteil besteht darin, dass bei dieser Form der Erschließung im Retrieval der abgefragte Zeitraum exakt mit den Zeitangaben im Zeitschlagwort der Titeldaten übereinstimmen muss. Weicht auch nur eine Ziffer ab, werden keine Treffer angezeigt, sofern eine reine Boolesche „UND“-Suche zugrunde liegt. In den meisten Fällen kann das Zeitschlagwort also nicht für die Recherche verwendet werden, sondern bringt nur bei der Anzeige einen Nutzen.

Beispiel:

Suche im Portal der DNB mit „Sozialgeschichte 1800–1988“:

Es wird nur ein Titel gefunden, der genau diese Jahreszahlerweiterung aufweist, obiger Titel „Sozialgeschichte Baden Württembergs“, der den Zeitraum bis 1989 umfasst, wird nicht gefunden:

Titel: Altes Handwerk im 20. Jahrhundert / Frieder Stöckle. – 1993

Schlagwörter: Stuttgart ; Handwerk ; Ländlicher Raum ; Sozialgeschichte 1800–1988.³¹

Es gibt theoretisch eine Reihe von denkbaren Varianten, um die maschinelle Auswertbarkeit der Zeitschlagwörter zu erhöhen und dadurch bessere Retrievalergebnisse zu erhalten. Eine Möglichkeit bestünde darin, das Zeitschlagwort in seiner gegenwärtigen Form (bei der die Jahreszahlen jeweils ganz exakt angegeben werden) durch normierte, gröbere Zeitabschnitte ähnlich den in den LCSH (Library of Congress Subject Headings³²) gebräuchlichen period subdivisions zu ersetzen. Eine andere Alternative wäre, die bisherige Praxis der Zeitangabe zwar beizubehalten, die freie Vergabe der Jahreszahlen jedoch durch das Anlegen von zeitlichen Normdatensätzen abzulösen und diese ihrerseits in eine Hierarchie von zeitlichen Normdatensätzen einzubinden. Bei jeder Erfassung eines neuen Zeitabschnitts wäre ein entsprechender zeitlicher Normdatensatz in der GND anzulegen, sofern der jeweilige Zeitraum nicht bereits in der exakt gleichen Form vorhanden ist. Beide Möglichkeiten verlangen allerdings zunächst einen sehr hohen Implementierungsaufwand.

29 Datensatz der Deutschen Nationalbibliothek, zuletzt geprüft am 28.03.2018, <http://d-nb.info/891225021>.

30 Für die Zeitschlagwörter ohne Jahreszahlerweiterung wird in der Masterarbeit empfohlen, unter bestimmten Bedingungen das Erscheinungsjahr bei dem Retrieval mitauszuwerten.

31 Datensatz der Deutschen Nationalbibliothek, <http://d-nb.info/552168610>, zuletzt geprüft am 28.03.2018.

32 Die LCSH bestehen aus Hauptschlagwörtern, die durch andere spezifiziert werden können; diese nennt man „subdivisions“. Es gibt eine ganze Reihe von „subdivisions“, unter anderem auch „period subdivisions“.

4. Modell „Abgleich von Zeitraumüberschneidung“

Bereits vor einigen Jahren hat Heidrun Wiesenmüller ein bisher unveröffentlichtes Modell entworfen,³³ welches die Zeitschlagwörter für das Retrieval auswertet und als „Abgleich von Zeitraumüberschneidung“ bezeichnet werden kann: Sie schlägt eine Methode vor, die auf einfachen mathematischen Abgleichen zwischen den Abfragezeiträumen der Nutzerinnen und Nutzer und den in den Titeldaten enthaltenen Jahreszahlangaben der Zeitschlagwörter beruht.³⁴ Sobald in irgendeiner Weise eine Überschneidung zwischen den beiden Zeitspannen Abfragezeitraum und Titeldaten vorliegt, werden die entsprechenden Katalogeinträge als Treffer ausgegeben (Abb. 3).

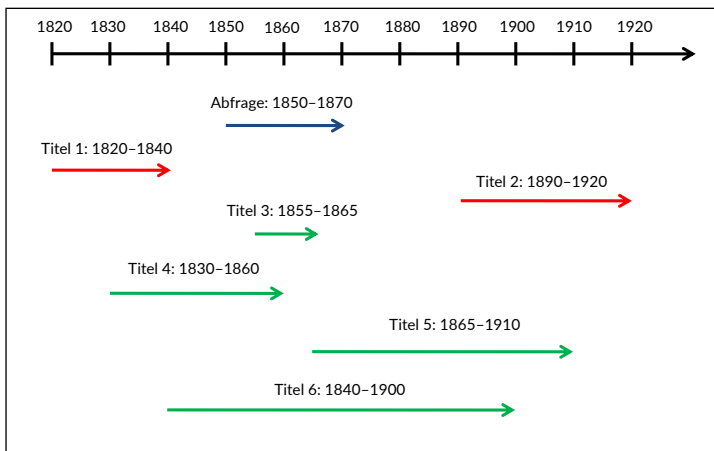


Abb. 3. „Abgleich von Zeitraumüberschneidung“, H. Wiesenmüller, 2011: Es wird die Zeitspanne 1850 bis 1870 abgefragt; Treffer sind alle Titel, deren erfasste Zeitspannen sich mit dem abgefragten Zeitraum um mindestens ein Kalenderjahr überlappen (Titel 3-6, grüne Zeitpfeile)

Dieses Verfahren bringt viele Vorteile mit sich, in erster Linie eine bessere Berücksichtigung aller für die Nutzerin oder den Nutzer tatsächlich relevanten Titel (Recall). Allerdings ergibt sich durch die Methode hinsichtlich des Anteils der Treffer mit Relevanz eine größer werdende Unschärfe, je länger die abgefragten Zeiträume werden. Denn wer Überblicksdarstellungen zu langen Epochen sucht, würde aufgrund des großen Schnittmengenbereichs des Abfragezeitraums viele irrelevante Titel in seiner Trefferliste erhalten, da sämtliche Titel, welche auch nur ein einziges Kalenderjahr an „Übereinstimmung“ mit dem Abfragezeitraum haben, automatisch in die Trefferliste aufgenommen werden, soweit die Suche nicht durch andere relevante Suchbegriffe eingeschränkt würde. Dieser

33 Die Ausführungen beziehen sich auf das unveröffentlichte Dokument Heidrun Wiesenmüllers. Heidrun Wiesenmüller, *Recherche über Zeitschlagwörter: Brainstorming*, unveröffentlichtes Manuskript (2011), 1-2.

34 Ohne dass damit der Aufwand des Erstellens eines Normdatensatzes einhergeht. Es würde ausreichen, die Jahreszahlen nicht direkt hinter das Zeitschlagwort zu schreiben, sondern in jeweils dafür vorgesehenen, eigenen Feldern abzuspeichern. Prinzipiell könnte das Erfassen des Zeitcodes (Paragraf 418, RSWK) als Vorlage hierzu dienen.

Ausfransung der Präzision in der Treffermenge könnte man mit einem sogenannten Toleranzfenster entgegenwirken.

5. Toleranzfenster-Tool für präziseres Suchen

Um sowohl einen möglichst vollständigen Recall als auch eine hohe Precision in der zeitlichen Suche zu erreichen, müsste man für das inhaltliche Zeitretrieval nach längeren Zeiträumen eine Möglichkeit vorsehen, die Anfangs- und Endjahre der zeitlichen Suche jeweils wiederum mit einer Art Toleranzfenster erweitern zu können, dessen Ausdehnung variabel einstellbar sein könnte. Alle Titel, deren Zeitschlagwörter mit Jahreszahlerweiterung keine Überschneidung mit den Zeiträumen der beiden Toleranzfenster aufweisen, gehen somit nicht mehr in die Trefferliste ein.

Zur Veranschaulichung des Sachverhalts ein Beispiel: Ein über einhundertfünfzig Jahre langer Abfragezeitraum (1789–1945) würde im Modell der Überschneidungsmengenauswertung viele und in der Mehrheit irrelevante Treffer hervorbringen, da sämtliche Titel, deren Zeitspanne komplett innerhalb oder auch nur zum Teil innerhalb dieser Epoche liegen, Eingang in die Trefferliste fänden. Meist möchten die Recherchierenden jedoch lediglich Titel angezeigt bekommen, die die von ihnen eingestellte Epoche in Gänze oder zumindest wesentliche Teile davon behandeln. Dazu sollen aber nach Möglichkeit auch Titel gehören, deren Anfangs- und Endjahre im Zeitschlagwort beispielsweise 1787–1955 oder 1792–1939 lauten, mithin also ähnliche, nur geringfügig abweichende Zeitparameter aufweisen.

Um diese Recherchewünsche möglichst exakt zu erfüllen, reicht die Methode der vorgestellten einfachen Überschneidungsmengenauswertung aufgrund der oben erläuterten Defizite nicht mehr aus. Sie wird daher um die Öffnung von Toleranzfenstern um die beiden Daten „1789“ und „1945“ herum ergänzt. Für das Anfangsjahr werden beispielsweise plus/minus fünf „Toleranzjahre“, für das Endjahr plus/minus zehn „Toleranzjahre“ eingestellt.³⁵ Alle Publikationen, deren Zeitschlagwortangaben keine Überschneidung mit den beiden Toleranzfensterzeiträumen aufweisen, fallen nun aus der Trefferliste heraus, was die Präzision des Rechercheergebnisses wesentlich erhöht. Umgekehrt können aber alle Titel, bei denen die Zeitschlagwortangabe der Anfangs- oder Endjahre innerhalb der Toleranzfenster liegt, in der Trefferliste erscheinen, da sie aufgrund ihrer großen Ähnlichkeit mit dem parametrisierten Suchzeitraum als relevant angesehen werden. Auch die beiden oben angeführten Titelzeiträume wären bei der beispielhaft skizzierten Ausdehnung der Toleranzfenster in der Treffermenge inkludiert.

In vielen Katalogen sind trunkierte Suchen möglich. In gewisser Weise entsprechen sie einem Toleranzfenster. Allerdings sind sie kaum skalierbar. Eine trunkierte Suche im Portal der DNB mit „Sozialgeschichte 1800–198*“ ergibt zehn relevante Treffer; die Abfrage „Sozialgeschichte 180*–198*“ bringt dagegen keine weiteren Suchergebnisse; wenn allerdings bereits nach der zweiten Stelle „Sozialgeschichte 18*–198*“ trunkiert wird, steigt die Trefferzahl rapide an, allerdings mit dem oben

³⁵ Das Toleranzfenster um das Anfangsjahr hat demnach die Ausdehnung 1784–1794, das Toleranzfenster um das Endjahr hat die Ausdehnung 1935–1955. Um das Jahr 1945 wird das Toleranzfenster deshalb weiter geöffnet, da somit sichergestellt werden kann, dass alle relevanten Publikationen zum Zweiten Weltkrieg und die unmittelbaren Nachkriegsjahre in der Treffermenge mitberücksichtigt werden.

beschriebenen Nachteil, dass die Präzision zurückgeht. Toleranzfenster bieten einen Mehrwert zur trunkierten Suche.

6. Fazit

Mithilfe der Verwendung von spezifischen Toleranzfenstern könnten also auch und insbesondere bei langen Suchspannen die beiden grundsätzlichen Anforderungen nach Recall und Precision beim zeitlichen Retrieval erfüllt werden: Die Trefferliste enthält alle Titel, die für die Nutzerin oder den Nutzer relevant sein können und spart dabei gleichzeitig sehr viele nichtrelevante Titel aus.

Es gibt in einigen Punkten der vorgestellten Variante „Abgleich von Zeitraumüberschneidung“ noch offene Fragen, beispielsweise wie eine nutzerfreundliche Realisierung des Toleranzfenster-Tools gewährleistet werden kann, oder auf welche Weise die zeitliche Parametrisierung am intuitivsten erfolgen sollte (über Zeitstrahl oder Eingabefenster). Sie müssen im konkreten Anwendungsfall unterschiedlich gehandhabt und auf die jeweiligen Anforderungen der betreffenden Informationseinrichtungen angepasst werden.

Die begonnene Thematik „Erfassung und Recherche der Zeitfacette“ stellt ein weites Feld dar. In einem späteren Artikel soll das Zeitretrieval außerhalb von Bibliothekskatalogen und die Anwendung der Time-Ontology auf die GND untersucht werden.

Literaturverzeichnis

- Deutsche Nationalbibliothek. *Regeln für die Schlagwortkatalogisierung*: RSWK. 4., vollständig überarbeitete Auflage 2017. Stand: März 2017. [urn:nbn:de:101-2017011305](https://nbn-resolving.org/urn:nbn:de:101-2017011305).
- Frommeyer, Jutta. *Zeitbegriffe und Zeitcodierungen in allgemeinbibliographischen Datenbanken: Eine Analyse von RSWK, LCSH und RAMEAU sowie Entwicklung eines theoretischen Modells für ein Zeitretrieval*. Berlin: Logos-Verl., 2003.
- IFLA. „Bibliographic Conceptual Models.“ Zuletzt geprüft am 28.03.2018. <https://www.ifla.org/node/2016>.
- IFLA. „Functional Requirements for Authority Data FRAD (2009).“ Zuletzt geprüft am 28.03.2018. <https://www.ifla.org/publications/functional-requirements-for-authority-data>.
- IFLA. „Functional Requirements for Bibliographic Records FRBR (1998).“ Zuletzt geprüft am 28.03.2018. <https://www.ifla.org/publications/functional-requirements-for-bibliographic-records>.
- IFLA. „Functional Requirements for Subject Authority Data FRSAD (2011).“ Zuletzt geprüft am 28.03.2018. <https://www.ifla.org/node/5849>.

- IFLA. „Library Reference Model (IFLA LRM) (2017).“ Zuletzt geprüft am 28.03.2018. <https://www.ifla.org/publications/node/11412>.
- International Organization for Standardization. „ISO 8601 Date and time format: Data elements and interchange formats – Information interchange – Representation of dates and times.“ Zuletzt geprüft am 28.03.2018. <https://www.iso.org/standard/40874.html>.
- Scheven, Esther. „Rezension zum Buch von Jutta Frommeyer: Zeitbegriffe und Zeitcodierungen in allgemeinbibliographischen Datenbanken.“ *Zeitschrift für Bibliothekswesen und Bibliographie* 50 (2003): 295–297. Zuletzt geprüft am 28.03.2018. http://zfbf.thulb.uni-jena.de/servlets/MCRFileNodeServlet/jportal_derivate_00001835/j03-h5-rez-1.pdf.
- Wening, Michael. *Die Zeitfacette in bibliothekarischen Erschließungen: Eine Analyse unter besonderer Berücksichtigung von sachlicher Erschließung und Retrieval*. Masterarbeit im Rahmen des Masterstudiengangs „Bibliotheks- und Informationswissenschaft“. Berlin: Humboldt-Universität, 2017. Zuletzt geprüft am 28.03.2018. <http://d-nb.info/1149163542>.
- Wiesenmüller, Heidrun. *Recherche über Zeitschlagwörter: Brainstorming*. Unveröffentlichtes Manuskript (2011).

Tagungsberichte

Bericht zum Workshop „Rechtliche Aspekte bei digitalen Forschungsdaten“ an der Europa-Universität Viadrina Frankfurt (Oder)

Im eindrucksvollen Logensaal der Europa-Universität Viadrina unweit des Grenzflusses Oder kamen am 30. Januar 2018 etwa 100 Personen zusammen, um sich über ein offenbar drängendes Themenfeld auszutauschen: Rechtliche Aspekte bei digitalen Forschungsdaten.¹ Das sehr divers zusammengesetzte Publikum zeigte dabei unterschiedliche Interessenlagen. Die Teilnehmenden mit einem eher juristischen Hintergrund verfolgten vor allem das Ziel, die konkreten Fragen und Rechtsaspekte der Realität des Forschungsdatenalltags entgegenzustellen und Lösungen für das Spannungsfeld zu finden bzw. zu diskutieren. Dagegen suchten Vertreterinnen und Vertreter aus dem Infrastrukturfeld rechtswissenschaftliche Expertise zu all den Fragen, die trotz intensiver Beschäftigung mit dem Thema bislang offenblieben. Die Forschenden schließlich hofften auf klare Auskünfte, wie sie das Forschungsdatenmanagement bzw. die Publikation von Forschungsdaten so gestalten können, dass sie nicht über eventuelle rechtliche Fallstricke stolpern.

Eine erschöpfende Einlösung dieser Erwartungen konnte naturgemäß nicht erwartet werden. Erfolgreich war die Veranstaltung allemal: Es ist gelungen, einige Problembereiche und Desiderate präziser einzukreisen und zu diskutieren. Und sie fand dafür am richtigen Ort statt. Die Viadrina ist mit ihren fachlichen Schwerpunkten der Wirtschafts-, Rechts- und Kulturwissenschaften vergleichsweise neu auf der Landkarte des Forschungsdatenmanagements zu verorten. An der Viadrina konnte der Bereich „Forschungsdaten & Forschungsinformation“ im Referat „Forschung und Wissenschaftlicher Nachwuchs“ etabliert werden. In einem europaweit einmaligen Schritt soll eine internationale Fakultät mit dem Leitthema der Digitalisierung und ihren wirtschaftlichen, rechtlichen und gesellschaftlichen Dimensionen gemeinsam mit der polnischen Adam-Mickiewicz-Universität in Poznan ins Leben gerufen werden, wie der Geschäftsführende Präsident der Hochschule, Stephan Kudert, in seiner Begrüßung erläuterte. So kann sich dort wohl in absehbarer Zeit ein datenrechtlicher, vielleicht sogar forschungsdatenrechtlicher Schwerpunkt etablieren. Die Veranstaltung, die gemeinsam von dem BMBF-Projekt FDMentor² und dem DFG-Projekt eDissPlus³ ausgerichtet wurde, lässt sich also auch als Auftakt verstehen. In der Eröffnungsrede von Hans-Gerd Happel, dem Direktor der Universitätsbibliothek an der Viadrina und Teilprojektleiter bei FDMentor, wurde dabei vor allem der Wunsch nach praxisorientierten Handlungsempfehlungen für Forschungseinrichtungen Ausdruck verliehen.

1 Vortragspräsentationen aller Vortragenden abrufbar unter http://www.forschungsdaten.org/index.php/Rechtliche_Aspekte_bei_digitalen_Forschungsdaten; siehe ferner auch im Viadrina-Logbuch den Eintrag „Der richtige Umgang mit digitalen Daten“ unter <https://www.europa-uni.de/de/struktur/unileitung/pressestelle/viadrina-logbuch/wissenschaft/20180130-Workshop-Forschungsdaten/Beitrag/index.html>.

2 <http://www.forschungsdaten.org/index.php/FDMentor>.

3 <https://www2.hu-berlin.de/edissplus/>.

Motivation eDissPlus

Aus Sicht von eDissPlus stellen die rechtlichen Unsicherheiten eine Hauptursache für die Zurückhaltung von Promovierenden beim Publizieren von Forschungsdaten dar. In den Befragungen wurde dieser Aspekt nahezu durchgehend herausgehoben. Interessant ist dabei, dass offenbar auch seitens der Gutachtenden selten verbindliche Auskünfte und Hinweise zur rechtskonformen Aufbereitung und Publikation von Forschungsdaten vermittelt werden können. Häufig wird das Problem an andere hochschulinterne Akteure delegiert - die Universitätsbibliothek, Forschungsdatenbeauftragte und vor allem die Rechtsstellen der Hochschule, die jedoch in den meisten Fällen die tatsächliche fachwissenschaftliche Dimension der Gemengelage schwer einschätzen können. Aus dieser Erkenntnis entstand bei eDissPlus der Wunsch, die damit verbundenen Fragestellungen zumindest in einer Weise zu systematisieren, die es ermöglicht, präzisere Hinweise an die Promovierenden beispielsweise über die im Projekt zu erstellenden Guidelines geben zu können. Der Workshop wurde für diesen Zweck zum Erfahrungsaustausch, zur Informationsgewinnung und in gewisser Weise auch zur eigenen Weiterbildung konzipiert.

Motivation FDMentor

In der Präsentation des als Mitveranstalter auftretenden Projekts eDissPlus wurde in Bezug auf die Studierenden festgestellt, was sich über weite Strecken auch über die Forschung insgesamt sagen lässt: Ohne rechtliche Unterstützung sind Forschende nicht dafür zu gewinnen, ihre Forschungsdaten bereitzustellen. Erfahrungen aus dem strategischen und operativen Forschungsdatenmanagement an vielen Einrichtungen unterstreichen dies, wie nicht zuletzt die Diskussionen bei der Workshop-tagung verdeutlicht haben. Für Dienste und Infrastrukturen erweist es sich daher als unerlässlich, eigene Expertise und Strukturen für einen professionellen Umgang mit Rechtsfragen zu entwickeln und konkret zu etablieren.

Wie das umgesetzt werden kann, wird vom BMBF-geförderten Projekt FDMentor untersucht. Es befasst sich mit der Frage, wie universitäre Strukturen erweitert und unter anderem auf Rechtsfragen bei Forschungsdaten zugeschnitten werden können. Im Unterschied zu anderen Projekten liegt der Fokus von FDMentor nicht auf Forschenden und der Auskunft zu einzelnen Rechtsfragen, sondern auf infrastrukturellen Veränderungen für eine bessere Beratung und Unterstützung der Wissenschaft. In der Praxis zeigt sich offen die Lücke, die FDMentor anspricht. Philipp Krahn unterstrich dies ebenfalls in seinem Vortrag am Beispiel des Datenschutzes: Aktuell wird die Verantwortung zur Einhaltung (datenschutz)rechtlicher Vorgaben meist allein auf die Forschenden selbst verlagert, die jedoch naturgemäß nicht in allen Fällen die notwendige juristische Beurteilungskompetenz aufweisen. Bislang zu wenig reflektiert sind die Haftungskonstellationen, vor allem aber auch deren Wirkung auf Forschende, wenn es um Nachnutzung von Forschungsdaten und um die Veröffentlichungsbereitschaft in Datenrepositorien geht.

Insgesamt ist die Lösung vermutlich weniger – so der Gesamteindruck aus den Diskussionen – Forschende juristisch in der erforderlichen Breite und Tiefe, d.h. sehr umfassend juristisch zu schulen,

sondern ihnen über den Auf- und Ausbau von Forschungsdatendiensten infrastrukturelle Kontaktstellen und Compliance-Workflows anzubieten, die ihre Arbeit juristisch flankieren und absichern.

Bei FDMentor soll weitergehend daran gearbeitet werden, welche wissenschaftsunterstützenden Zentralstellen für Rechtsfragen zuständig und entsprechend qualifiziert werden sollten. Dieser Ansatz wurde auf dem Workshop vorgestellt und diskutiert. Inwieweit dieser Ansatz realistisch ist, kann gut eingeschätzt, wer an den Universitäten täglich mit Forschungsdaten umgeht: Hauptsächlich zum Workshop eingeladen waren daher Forschungsdatenreferentinnen und Forschungsdatenreferenten sowie entsprechend befasstes Fachpersonal aus den Bibliotheken, Justizariaten und Rechenzentren an Wissenschaftseinrichtungen.

Keynotes

Ursprünglich als reiner Workshop angedacht, erforderte die unerwartet hohe Nachfrage eine methodische Anpassung der Veranstaltung. So wurde am Vormittag zunächst ein Vortragsblock vorgesehen, der in das Themenfeld aus unterschiedlichen Perspektiven einführte. Die Keynotes umfassten dabei die Vorstellung der Projekte FDMentor und eDissPlus, die zur rahmengebenden Infrastruktur forschen, das Thema der Herausforderungen der Forschungspraxis, die am Beispiel der kulturwissenschaftlichen Fakultät der Viadrina vorgestellt wurden, sowie die Beleuchtung des Themas aus fachjuristischer Perspektive am Beispiel des Datenschutzes.

Die infrastrukturelle Perspektive

Im Projekt FDMentor sollen für die Nachnutzung geeignete Roadmaps, Handlungsempfehlungen und Good-Practice-Beispiele zur strategischen Entwicklung und Verbesserung des Forschungsdatenmanagements an deutschsprachigen Hochschulen erarbeitet werden. Schwerpunkt ist das Bereitstellen von Werkzeugen für eine Strategieentwicklung zum Forschungsdatenmanagement, von Modellen für institutionelle Forschungsdaten-Policies sowie eines Beratungs- und Trainingskonzepts. Partner des Verbundprojektes sind die Humboldt-Universität zu Berlin, die Freie Universität Berlin, die Technische Universität Berlin, die Universität Potsdam und die Europa-Universität Viadrina Frankfurt (Oder). Dabei sollen vor allem die bestehende Expertise an den Zentraleinrichtungen der Verbundpartner einbezogen und generalisierbare Strategien abgeleitet werden. Da sich das Projekt noch in einer frühen Phase befindet, bildete die Keynote vor allem die Fragestellungen und den Realisierungsansatz ab.

Kennzeichnend für die aktuelle Wissenschaftslandschaft ist, dass es nicht nur darum geht, Forschungsdaten zeitgemäß und rechtssicher zu organisieren. Ergänzend kommt der Anspruch von Open Science oder Open Scholarship als weitere Herausforderung hinzu.

Viele Infrastruktureinrichtungen wie Universitätsbibliotheken, Rechenzentren, Repositorienbetreiber oder Fachdienste sehen sich daher mit dem Anspruch konfrontiert, eine politisch oder von Förderern und aus der Community geforderte offene Wissenschaft durch passende Angebote zur Veröffentlichung von Forschungsdaten zu unterstützen. Das zweijährige DFG-Projekt eDissPlus, eine Kooperation der Humboldt-Universität zu Berlin und der Deutschen Nationalbibliothek (DNB), versucht diese

Entwicklung mit dem Fokus auf dissertationsbezogenen Forschungsdaten zu begleiten. Da es sich kurz vor dem Abschluss befindet, konnten eine Reihe von Einsichten präsentiert werden, die an vielen Stellen verdeutlichen, wie komplex die Herausforderungen und wie selten Lösungen vorhanden sind.

Promotionsvorhaben sind ein interessanter Anwendungsfall, da sie einerseits vom Zweck her formal restringiert sind, andererseits die gesamte Forschungsvielfalt abbilden, da es sich um Forschungsdaten aus sämtlichen Fachbereichen handelt, die dementsprechend in unterschiedlichsten Formen und Formaten auftreten können. Das alle Disziplinen (zumindest an der Humboldt-Universität zu Berlin) verbindende Element und zugleich ein zentrales Hindernis für Forschungsdatenpublikationen ist, dass diese Ergänzung zur Dissertation im Rahmen von Promotionen kaum explizit vorgesehen ist und entsprechend auch keine Berücksichtigung in den Prüfungsordnungen findet. Die rechtliche Unsicherheit beginnt also schon an dieser Stelle. Technisch ist die Veröffentlichung von dissertationsbezogenen Forschungsdaten dennoch möglich. Das Projektziel der Entwicklung einer entsprechenden Lösung sowie von Workflows für die Pflichtablieferung solcher Daten an die Deutsche Nationalbibliothek im Rahmen ihres Sammelauftrages wurden erfolgreich realisiert.⁴ Ebenso ist die notwendige Begleitung über Beratung und Guidelines an der Humboldt-Universität vorbereitet worden und wird derzeit implementiert.

In Übereinstimmung mit anderen Studien ermittelte eDissPlus, dass zwar viele Promovierende grundsätzlich bereit sind, ihre Forschungsdaten zu teilen, aber nur wenige dies tatsächlich tun. Neben der Unsicherheit scheinen dabei hauptsächlich zwei weitere, miteinander in Beziehung stehende Faktoren entscheidend zu sein: Erstens gewinnen Promovierende im unmittelbaren Promotionszusammenhang nichts, wenn sie ihre Forschungsdaten bereitstellen. Zweitens wird der für eine ordentliche Forschungsdatenpublikation anfallende erhebliche Aufbereitungsaufwand der Daten bislang kaum bei der Zeit- und Umsetzungsplanung des Promotionsvorhabens berücksichtigt. In der Praxis bedeutet eine dissertationsbegleitende Forschungsdatenpublikation für die meisten Befragten eine nicht notwendige aber sehr aufwändige Zusatzbelastung, von der man im Zweifel lieber absieht. Viele Promovierende wollen zudem die Kontrolle über die Weitergabe eigener Forschungsdaten behalten und selbst bestimmen, wem sie welche Daten unter welchen Bedingungen zur Verfügung stellen.

Auf die Frage nach gewünschter Unterstützung geben die Promovierenden neben einer verlässlichen Archivierung vor allem eine Beratung bei rechtlichen Fragen an (z.B. Zugangseinschränkungen, Umgang mit sensiblen Daten, Nutzung von Lizenzen, Haftungsrisiken, Dateneigentum). Tatsächlich handelt es sich bei der bestehenden Rechtsunsicherheit, insbesondere in Bereichen der empirischen Sozialforschung, um eine der größten Hürden für das Publizieren von Forschungsdaten. Die zentrale Herausforderung aus Sicht der Beratung besteht schließlich in dem Spannungsverhältnis, dass Guidelines für Forschungsdatenpublikationen vergleichsweise allgemeine Richtlinien bzw. Empfehlungen angeben müssen, obwohl oftmals die Prüfung des Einzelfalles unumgänglich erscheint.

4 Siehe dazu den Beitrag von Dirk Weisbrod in diesem Heft (S. 72–78, <https://doi.org/10.5282/o-bib/2018H2S72-78>), sowie Weisbrod, Dirk, „Pflichtablieferung von Forschungsdaten,“ *Dialog mit Bibliotheken* 30, Nr. 1 (2018): 26–27, <https://d-nb.info/115432320X/34>.

Die fachwissenschaftliche Perspektive

Wie komplex sich das Thema digitale Forschungsdaten in der konkreten Wissenschaftspraxis entfalten kann, legten Dorothea Horst und Silva Ladewig vom Lehrstuhl für Sprachgebrauch und multimodale Kommunikation an der Europa-Universität Viadrina Frankfurt (Oder) in ihrer Präsentation überzeugend dar. In ihrer Forschung kommen Multimodalität und Multimedialität und die damit verbundenen Daten in einer Vielfalt zusammen, die sich kaum mit generalisierenden Rechtsvorschriften einfangen lässt. Wissenschaftlerinnen und Wissenschaftler sind sich dessen bewusst und schildern, wie berichtet wurde, regelmäßig, wie sie Nutzungen für konkrete Forschungs- und Lehrzwecke rechtskonform realisieren wollen, sich dabei mitunter aber vermutlich in rechtlichen Grauzonen bewegen.

Die beiden Referentinnen erläuterten, dass sich für zahlreiche Aspekte sehr pragmatische und aus der Praxis entstehende Lösungen finden, die sich als fachkulturell tauglich, aber nicht immer zufriedenstellend erweisen. So wird die Anonymisierung bei mimischen Informationen dadurch gewährleistet, dass fotografierte Gesichter in Zeichnungen umgewandelt werden. Naturgemäß gehen dabei jedoch Informationsanteile verloren. Insofern bewegen die Forschenden sich ständig im Spannungsfeld datenschutzrechtlicher Normen und fachwissenschaftlicher Ziele, nämlich der Zeigbarkeit und Weiternutzbarkeit, und stoßen regelmäßig an die Grenzen der rechtskonform realisierbaren Zugänglichmachung. Deutlich wurde zugleich, dass sich ein interdisziplinärer Dialog zu forschungspraktischen Erfahrungen und Lösungen anbietet. Abschließend bestätigten die Forscherinnen die Notwendigkeit von Anlaufstellen bei Rechtsfragen und Beratungsangeboten, wie etwa universitären Datenschutzbeauftragten oder Forschungsreferaten.

Die juristische Perspektive

Philipp Krahn von der Technischen Universität Dresden konfrontierte das Publikum mit der geballten Komplexität des Datenschutzrechtes, was bei vielen der Teilnehmenden sichtlich zu einem Gefühl der Überwältigung führte. Krahn ist wissenschaftlicher Mitarbeiter im Projekt DataJus, das in derselben Förderlinie wie FDMentor vom BMBF gefördert wird und sich auf die rechtswissenschaftliche Dimension des Phänomens Forschungsdaten konzentriert.

Die erweist sich häufig als sehr komplex, nicht zuletzt da mindestens drei unterschiedliche gesetzliche Rahmenbestimmungen zu beachten sind: die Datenschutzgesetze der Bundesländer, des Bundes sowie die ab Mai 2018 geltende EU-Datenschutz-Grundverordnung. Zusätzlich sind Leitlinien für eine gute wissenschaftliche Praxis, Ethikrichtlinien sowie institutionelle bzw. fachkulturelle Data Policies zu beachten. Alle diese Regelungen streben idealerweise nach wechselseitiger Harmonie, erreichen diese aber nicht in jedem Fall. Die informierte Einwilligung, die es bei der Forschung mit personenbezogenen Daten in der Regel einzuholen gilt, ist dann eine sinnvolle Lösung, wenn sie alle eventuellen Nutzungsmöglichkeiten, also im Prinzip auch zum Zeitpunkt der Einwilligung unbekanntes Nutzungszwecke mit erfasst. Zugleich entfaltet sie angesichts der komplizierten Gemengelage teilweise eine Komplexität, die die zu Informierenden möglicherweise überfordert, so dass im Ergebnis die Frage gestellt werden muss, ob die Einwilligung tatsächlich zureichend „informiert“ geschehen kann. Dass es obendrein zahlreiche Forschungsszenarien gibt, bei denen die formale Vorab Einwilligung

die Forschung selbst direkt unterlaufen kann, sollte Sabine Imeri vom „Fachinformationsdienst Sozial- und Kulturanthropologie“ in ihrem Vortrag anschaulich darlegen. Das von Philipp Krahn vorgeschlagene Phasenmodell zu den datenschutzrechtlichen Vorgaben für Forschungsdaten dürfte dessen ungeachtet als strukturierte Handreichung dennoch bei allen Infrastrukturverantwortlichen sehr willkommen sein.

Impulsvorträge und Arbeitsgruppen

Am Nachmittag wurden in einem zweiten Block einzelne Aspekte zunächst mit Impulsvorträgen und im Anschluss in sechs Arbeitsgruppen vertieft. Auch hier bildeten die Dimensionen Fachwissenschaft, Infrastruktur und Recht die thematische Einfassung.

Sabine Imeri: Ethnologische Forschung

Dr. Sabine Imeri ist einerseits Ethnologin und ehemalige wissenschaftliche Mitarbeiterin und Lehrbeauftragte am ethnologischen Institut der Humboldt-Universität zu Berlin und andererseits Mitarbeiterin im „Fachinformationsdienst Sozial- und Kulturanthropologie“ an der Universitätsbibliothek der Humboldt-Universität mit dem Schwerpunkt Forschungsdatenmanagement. Insofern ist sie mit den zum Teil konfligierenden Perspektiven aus fachwissenschaftlicher Forschung und Infrastrukturentwicklung bestens vertraut. In ihrem Impulsvortrag stellte sie die Vielfalt ethnologischer Forschungsdaten und die Bedeutung von Kontextinformationen für die zumeist qualitative empirische Sozialforschung heraus. Diese umfasst im besonderen Maße auch personenbezogene und schutzbedürftige Daten wie Angaben bezüglich Herkunft, sozialem Status oder Religionszugehörigkeit. In ihrer Arbeitsgruppe wurde deshalb die Bedeutung von zugangsbeschränkten Forschungsdatenarchiven herausgestellt, die es ermöglichen Forschungsdaten unter rechtskonformen Bedingungen zu archivieren und nur soweit wie möglich zugänglich zu machen.

Ellen Euler: Unendliche Weiten: Research/Data/Rights Management?

Ellen Euler, Professorin am Fachbereich Informationswissenschaft der Fachhochschule Potsdam, stellte den Forschungsdatenprozess von der Projektplanung bis zur Datenarchivierung dar. Dabei sind drei Managementbereiche bei Daten zu beachten: Erstens die technische, zweitens die dokumentarische und drittens die rechtliche und ethische Dimension.

Wie aktuelle Forschungsarbeiten an der Fachhochschule Potsdam verdeutlichen, werden bei beratenden Forschungsdatendiensten rechtliche Fragestellungen besonders stark nachgefragt. In der Arbeitsgruppe wurde daran anschließend diskutiert, welche bedarfsorientierten Qualifizierungsangebote entwickelt werden könnten, etwa Weiterbildungsmodule zum Data Rights Manager.

Ben Kaden: Interessenskonflikte bei der Publikation von Forschungsdaten

Für das Projekt eDissPlus, dessen Erkenntnisse hinter diesem Schwerpunkt standen, fokussierte Ben Kaden in seiner Präsentation den sehr spezifischen, aber aus Sicht der Hochschulinfrastrukturen sehr wichtigen Aspekt der Publikation von Forschungsdaten im Zusammenhang mit Dissertationen. Diese wird allgemein als sehr erstrebenswert bewertet, praktisch aber aus verschiedenen Gründen nicht umgesetzt. Infrastrukturen stehen daher vor der Herausforderung, dass sie gern passende Dienste

entwickeln möchten, aber dafür kaum auf Erfahrungswerte zurückgreifen können. Die Arbeitsgruppe diskutierte konkret rechtliche Konfliktlagen, die einer freien Publikation von dissertationsbegleitenden Forschungsdaten oft im Weg stehen.

Neben dem großen Thema des richtigen Umgangs mit personenbezogenen Daten, der eine Nachnutzung von Forschungsdaten grundlegend ausbremst, lassen sich zwei weitere zentrale Problemfelder ausmachen. Auf der Seite des Zugangs sind Forschungsdaten Dritter, also Archive und Behörden, oft restriktiven Hausregelungen unterworfen, die eine Nutzung stark reglementieren und eine Weitergabe oft verhindern. Solange es an dieser Stelle keine umfassende Lösung gibt, können Promovierende eventuellen Ansprüchen an eine Publikation dieser Daten nicht rechtssicher nachkommen. Die zweite Hürde liegt in der Unsicherheit hinsichtlich urheberrechtlicher Ansprüche, die sich aus der Erhebung oder Ver- bzw. Bearbeitung von Forschungsdaten und -materialien für die Forschenden ergeben. Dies wirkt sich prinzipiell auf die Bereitschaft der Datenpublikation und auf Unsicherheiten bei der Lizenzzuweisung aus. Hier ist vor allem Beratung notwendig. Gleiches gilt für eventuelle verlagsvertragliche Regelungen zur Forschungsdatenpublikation.

Katharina de la Durantaye: Urheberrechte an Forschungsdaten –ein Überblick

Zunächst gab Katharina de la Durantaye, Professorin für Bürgerliches Recht an der Humboldt-Universität zu Berlin, eine Einführung in das Urheberrecht, in der sie vor allem das Kriterium der Werkhöhe betonte. Daneben ging sie auf Lichtbildwerke und Lichtbilder ein sowie auf den rechtlichen Schutz von Datenbanken. Im Resümee zur Arbeitsphase formulierte sie eine zumindest für die Perspektive Urheberrecht etwas Linderung versprechende Einsicht: In den meisten Fällen dürften die urheberrechtlichen Unsicherheiten nicht mit der realen Gefahr einer Rechtsverfolgung verbunden sein. Im Bereich des wissenschaftlichen Publizierens von digitalen Forschungsdaten sind bislang kaum Klagen an Gerichten eingereicht worden. Das heißt aber auch, dass es kaum Gerichtsurteile gibt, die Präzedenzfälle zur Orientierung darstellen.

Thomas Hartmann: Die Datenschutzbeauftragten: Inneruniversitärer Ansprechpartner für Datenschutzfragen rund um digitale Forschungsdaten?

Nach einem Impulsvortrag diskutierte Thomas Hartmann in einer Arbeitsgruppe, inwieweit universitäre Datenschutzbeauftragte jeweils die Universitätsangehörigen beim Umgang mit digitalen Forschungsdaten beraten, begleiten, informieren, schulen und sonstig unterstützen können. Ausgangspunkt dafür ist die gesetzlich vorgegebene Position der (universitären) Datenschutzbeauftragten, wie sie ab Mai 2018 von der EU-Datenschutz-Grundverordnung festgelegt ist. Bei einem Erfahrungsaustausch erörterten die Teilnehmenden, welche Verankerung, Ressourcen und Tätigkeitsschwerpunkte die Datenschutzbeauftragten bislang an Universitäten haben und welche Chancen, aber auch welche Grenzen sich daraus für die Datenschutzerfordernisse bei Forschungsdaten ergeben könnten.

Philipp Krahn: Datenschutzrecht

Philipp Krahn diskutierte seinen Vortrag vom Vormittag mit den Teilnehmenden in einer Arbeitsgruppe, woraus sich eine Reihe von Vertiefungsfragen ergaben, die in der Forschungspraxis regelmäßig auftauchen, für die es aber (noch) kaum Lösungen gibt. Die Haftung spielt dabei eine wichtige Rolle, aber ebenso, wie der Umgang mit personenbezogenen Daten in Archivalien datenschutzrechtlich

bewertet werden sollte, welche neue Formen einer De-Anonymisierung die Big-Data-Analytik möglicherweise hervorbringt sowie die Frage nach der datenschutzrechtlichen Abstimmung bei internationalen Kooperationen.

Ausblick

Die strukturierte Behandlung der Rechtsfragen rund um Forschungsdaten ist ein bislang wenig beschrittenes Terrain. Das gilt für die infrastrukturelle ebenso wie für die fachjuristische Betrachtung. Mit einer Überblicksgrafik „Terra incognita - digitale Forschungsdaten auf der Suche nach einer rechtlichen Heimat“ umriss FDMentor-Projektmitarbeiter Thomas Hartmann zum Workshopende die zahlreichen, vielfältigen Rechtsgebiete, die bisher allesamt nicht wissenschaftsnah auf digitale Forschungsdaten und ihre Nutzung angepasst sind. FDMentor priorisiert bestimmte rechtliche Handlungsfelder und wird prüfen, inwieweit Verantwortlichkeiten und Expertisen der jeweils zuständigen Zentralstellen an den Universitäten in das Forschungsdatenmanagement einbezogen werden können.

So könnten seit langem vorhandene, akzeptierte Strukturen an Universitäten konkret aktiviert werden, um die hohen Hürden einer oft wenig wissenschaftsfreundlichen bzw. wissenschaftspraktikablen Gesetzeslage ein Stück weit zu reduzieren sowie die kaum eindeutig festgezurrten Gesetzesvorgaben verantwortungsbewusst zu interpretieren. Im universitären Gang der Forschung geben ohnehin oft tradierte fachkulturelle Ansprüche und Gepflogenheiten, im Idealfall auch wissenschaftsethische Erwägungen, die eigentlichen Leitlinien vor. Diese müssen zwar prinzipiell im Einklang mit den teils vielschichtigen gesetzlichen Bestimmungen liegen, aber forschungsdatenspezifische Entscheidungen können überall dort, wo ein Interpretationsspielraum vorliegt, offensiver in Richtung forschungsnaher und praktikabler Handhabung gefasst und getroffen werden. Dahingehend richtete sich auch der zuerst benannte Wunsch von Teilnehmenden bei der abschließenden Plenardiskussion: Welche Foren, Stellen und Netzwerke gibt es für die weiterführende Diskussion von Rechtsfragen? Vereinzelt waren in den Arbeitsphasen disziplinär ausgerichtete, nationale Kompetenzzentren angesprochen worden. Wie können Universitäten die Rechtsfragen zu den Forschungsdaten abbilden, aufgreifen und forschungsfreundlich auflösen? Diese Herausforderung wird das Projekt FDMentor bis zum Projektende Anfang 2019 begleiten.

Die Workshoptagung zeigte, dass es in vielen Anwendungsfeldern aus unterschiedlichen Gründen kaum möglich ist, alle denkbaren Eventualitäten juristischer Unschärfe zu berücksichtigen, ohne wissenschaftliche Erkenntnisinteressen zu gefährden. Die Aufklärung und Weiterbildung der Forschenden selbst dürfte hier nur als ergänzendes Mittel wirken und hat an vielen Stellen ihre Grenzen. Da fast unvermeidlich das Prinzip der Wissenschaftsfreiheit mit anderen Rechtsprinzipien potentiell in Konflikt geraten kann, setzt dies eine Doppelkompetenz bei denjenigen voraus, die innerhalb der Einrichtungen für die Forschungsdaten(rechts)beratung zuständig sind. Sie müssen die Fachkultur und Forschungsstruktur ebenso kennen wie die maßgeblichen rechtlichen Rahmenbedingungen. Auch die aktuell höchstqualifizierten Akteure in diesen Rollen müssen einräumen, dass genau dieses Qualifikationsprofil bisher nicht systematisch vermittelt wird.

Für die als Veranstalter auftretenden Projekte FDMentor und eDissPlus ergeben sich aus dem Workshop ähnliche, aber strukturgemäß anders gelagerte Erkenntnisse. Das Projekt eDissPlus befindet sich in seiner Abschlussphase, die eine Workshop-Road-Show beinhaltet, für deren Ausgestaltung wichtige weitere Gesichtspunkte erhoben werden konnten. Für FDMentor, das noch am Anfang steht und als Verbundprojekt eine Verlängerungsperspektive anstrebt, zeigt sich, dass nach wie vor ein sehr grundständiger Bedarf zur Systematisierung, Aufbereitung und Vermittlung des Themenfelds Forschungsdatenmanagement und -publikation, auch über die rechtlichen Fragen hinaus, besteht.

Thomas Hartmann, Europa-Universität Viadrina Frankfurt (Oder)

Ben Kaden, Humboldt-Universität zu Berlin

Michael Kleineberg, Humboldt-Universität zu Berlin

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S193-201>

Konferenz und Erfahrungsaustausch

26 Fachreferentinnen und Fachreferenten geisteswissenschaftlicher Fächer treffen sich in Bensberg (ZBIW-Seminar 26.–27.02.2018)

Was als informelles Treffen in Nordrhein-Westfalen vor sieben Jahren begann, hat sich zu einer jährlichen nationalen Fortbildungsveranstaltung entwickelt: Seit 2011 treffen sich Fachreferentinnen und Fachreferenten geisteswissenschaftlicher Fächer zur selbstorganisierten Fortbildung. Der diesjährige Austausch fand zum ersten Mal¹ in zweitägiger Klausur statt. Aus den Anregungen der letzten beiden Treffen, die noch jeweils für die Fachreferate Geschichte und Philologien getrennt durchgeführt wurden, und aus aktuellen Entwicklungen hatte das fünfköpfige Organisationsteam ein fachlich und inhaltlich gemischtes Programm zusammengestellt. Die bewährte Mischung der Beiträge von externen Vortragenden und aus den eigenen Reihen wurde beibehalten, die Moderation wieder von den Organisatorinnen übernommen. In den Blöcken Bestandsmanagement, Open Access in den Geisteswissenschaften, Wissenschaft und Wissenschaftsvermittlung sowie Wissenschaftskommunikation und Services für die Wissenschaft wechselten sich Vorträge und Plenarrunden ab. So entstand genug Raum für eine gelungene Mischung aus Anregung, Wissen, Austausch- und Diskussionsgelegenheit.

Bestandsfragen: Überlegungen zu Aussonderungen und Forschungsdatenmanagement

Am ersten Vormittag unter der Leitung von Dr. Alice Rabeler (ULB Bonn) sorgte das Referat „Aussonderung strategisch und praktisch“ von Dr. Ute Olliges-Wieczorek (ULB Düsseldorf) für eine die anschließende Diskussion bereichernde Mischung aus Theorie und Praxis: Ausgelöst durch den drohenden Wegfall von Archivflächen an der Universität Düsseldorf entstand die Notwendigkeit, über Aussonderungspotential nachzudenken. Aus dem gesetzlich verankerten Sammelauftrag als Landesbibliothek ergeben sich praktische Kriterien für Bestände, die unbedingt bewahrt werden müssen. In Düsseldorf wurden für die sonstigen Bestände Bewertungskriterien entwickelt, die als Ausgangspunkt für eine NRW-weit umsetzbare Strategie zur kooperativen Aussonderung dienen. Kernelement ist dabei die Kennzeichnung von Beständen – für Zeitschriften in der ZDB, für Monografien im HBZ-Verbundkatalog – für deren Archivierung sich eine Bibliothek zuständig und bereit erklärt. Dieser Vorschlag soll Anstoß geben für eine Entschließung in der Arbeitsgemeinschaft der NRW-Universitätsbibliotheksleitungen, um mit einem regionalen Konzept gerüstet zu sein, bevor lokal die Notwendigkeit auftaucht größere Aussonderungen realisieren zu müssen. Die Diskussion im Plenum ergab, dass die wenigsten Bibliotheken in NRW bisher dezidierte Archivierungskonzepte für ihre einzelnen Bestandsgruppen verfasst haben. Aus der Nutzungsanalyse magazinerter Zeitschriftenbände ergibt sich die dringende Erfordernis, solche nur im Print verfügbaren Texte über digitalisierte Inhaltsverzeichnisse besser zugänglich zu machen. Der zunehmende Bedarf an

1 Die jährlichen Treffen finden seit 2011 statt. S.a. Karolin Bubke und Dorothee Graf, „Fortbildung für Fachreferenten von WBs in Bielefeld und Bonn: von Kooperationsmöglichkeiten über Discovery Systeme zu E-Humanities - Fachreferat Moderne Philologien,“ *ProLibris* 20, Nr. 2 (2015): 88–89; sowie Yvonne Brzoska und Rosemarie Kosche, „ULB Düsseldorf – Erfahrungsaustausch für die Fachreferate Philologien sowie Geschichte und Kunstgeschichte,“ *ProLibris* 22 Nr. 3 (2017): 140.

Aussonderungen sei nur eine Frage der Zeit, denn die DBS-Zahlen zeigen einen steten Zuwachs an Medien in allen Bibliotheken, bei gleichzeitig steigenden Aussonderungsaktivitäten.

Während derartiges Bestandsmanagement großer Print-Bestände unter dem Vorzeichen der Digitalisierung praktiziert wird, erweitert sich der bibliothekarische Blick auf Forschungsdaten, die auch in den Geisteswissenschaften nicht unbekannt sind, auch wenn sie dort bislang meist Quellen heißen.² Auf nationaler Ebene wird eine Strategie zum Forschungsdatenmanagement diskutiert und entwickelt: die nationale Forschungsdateninfrastruktur (NFDI). Um gezielt auch die nordrhein-westfälischen Universitäten und Fachhochschulen einzubeziehen und zu adressieren, gibt es eine Landesinitiative NFDI der Digitalen Hochschule NRW. Um diese zu befördern und die Kartierung der einzelnen Strategiedebatten in NRW zu befördern, stellten die beiden an der UB Duisburg-Essen angesiedelten Projektmitarbeiterinnen Dr. Nina Winter und Marleen Grasse die Landesinitiative kurz vor und nutzten die Mittagspause, um sich weiter zu vernetzen.

Open Access in den Geisteswissenschaften

Der von Dorothee Graf, UB Duisburg-Essen, moderierte Montagnachmittag war ganz dem Thema Open Access gewidmet mit speziellem Fokus auf die Geisteswissenschaften und deren Publikationskultur.

Mit Language Science Press³ wurde „ein Community-driven Open-Access-Verlag im Übergang vom Projekt zum tragfähigen Geschäftsmodell“ vorgestellt.⁴ Sebastian Nordhoff schilderte Stationen, Grundsätze und Erfolgskriterien der noch recht kurzen Verlagsgeschichte und skizzierte die derzeitige Kooperation mit Knowledge Unlatched,⁵ um die Finanzierung der zukünftigen Publikationen zu sichern: Durch Pledges von je 1.000€ jährlich für die nächsten drei Jahre sichern die beteiligten Bibliotheken kooperativ die Betriebskosten des Verlags (rund 100.000€/Jahr) und ermöglichen 30-50 Publikationen im echten goldenen Open Access. Inhaltlich bedient der Verlag verschiedenste Teilbereiche aus der allgemeinen, sprachspezifischen und vergleichenden Linguistik. Die Titel erscheinen in zahlreichen Herausgeber-geführten Reihen und bedienen somit eine Community von weltweit rund 25.000 Sprachwissenschaftler/innen. Transparenz von Kosten, Review und Nutzung sind Kernelement der Verlagspolitik, Kennzahlen werden auf den Webseiten sowie im Blog⁶ veröffentlicht.

Die Förderung von Open Access ist in den Niederlanden auf ganz andere Art möglich und strategisch aufgestellt, wie der Althistoriker Dirk van Gorp, Open Access Officer der Universitätsbibliothek

2 So Peter Andorfer, *Forschungsdaten in den digitalen Geisteswissenschaften: Versuch einer Konkretisierung*. DARIAH-DE Working Papers Nr. 14 (Göttingen: DARIAH-DE, 2015), 13. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-7-2>.

3 Language Science Press, zuletzt geprüft am 27.04.2018, <http://langsci-press.org/>.

4 Vortragsfolien veröffentlicht unter <https://github.com/langsci/lsp-presentations/blob/master/2018dhd/presentation.pdf>, zuletzt geprüft am 27.04.2018.

5 S. <http://www.knowledgeunlatched.org/>, zuletzt geprüft am 27.04.2018. KU (Knowledge Unlatched) ist ein non-profit-Dienstleister für Bibliotheken und Verlage, entstanden 2012 aus dem Zusammenschluss verschiedener Akteure aus dem Verlags-, Publikations- und Bibliothekswesen. Ziel der GmbH ist es, wissenschaftliche Texte im Open Access verfügbar zu machen durch konsortiale Finanzierung mehrerer Bibliotheken.

6 S. <https://userblogs.fu-berlin.de/langsci-press/>, zuletzt geprüft am 27.04.2018.

Nijmegen, erläuterte: „Pushing and pulling the humanities towards open access journals publications: OLH, LingOA and other efforts at Radboud University Library, Nijmegen, and the Netherlands“. Durch sog. Deals auf nationaler Ebene mit einzelnen Verlagen wird Wissenschaftlerinnen und Wissenschaftlern aller niederländischen Universitäten ermöglicht, mit hohen Rabatten in anerkannten OA-Zeitschriften zu publizieren. Dabei sind die Geisteswissenschaften zwar ebenso zögerlich wie in Deutschland und bislang noch überwiegend an Buch(beitrags)publikationen interessiert, doch bietet die Open Library of the Humanities⁷ eine Plattform, um bekannte und neue Zeitschriften zu hosten, sofern sich die Herausgeber vom bisherigen Verlag lösen können. Da der Leitsatz „own the brands“ nicht von allen Herausgebern oder Verlagen eingehalten wird, wurde im prominentesten Beispiel der Titel *Lingua* durch den Umzug von Elsevier auf OLH geändert zu *Glossa*. Zahlreiche niederländische Universitäten haben sich zudem über die Stiftung LingOA⁸ zusammengeschlossen, um den Übergang von bislang subscriptions-finanzierten Zeitschriften auf ein „fair open access publishing model“ zu begleiten und für die nächsten fünf Jahre zu finanzieren.

Das abschließende Referat war der „Erfassung und Förderung von OA-Publikationen aus der Uni Duisburg-Essen“ gewidmet. Der Praxisbericht von Katrin Falkenstein-Feldhoff, OA-Beauftragte der UB, zeigte mit Zahlenmaterial aus der Hochschulbibliografie und dem Publikationsfonds der Universität Duisburg-Essen deutlich, wie stark bestimmte Fächer und Verlage von der inzwischen gängigen Förderung von APCs profitieren.⁹ Da dabei die Geisteswissenschaften fast völlig unberücksichtigt bleiben, entstand ein Kooperationsprojekt mit Verlagen, um im Rahmen des bestehenden Publikationsmodells hybride Publikationen zu fördern und Nutzungszahlen zu erheben. In parallelen Informationsmaßnahmen wie Schulungen, Handreichungen oder Vorträgen wird das Bewusstsein der Forschenden und Publizierenden für Open Access gesteigert und über rechtliche Möglichkeiten sowie Kosten und Nutzen von Open Access in den Geistes- und Sozialwissenschaften aufgeklärt. Damit soll die Diskussion um transparente und finanzierbare Modelle zwischen allen Akteuren des Publikationsprozesses befördert werden. Dieses Projekt wird vom BMBF unter dem Namen OGe-SoMo, Förderung von Open-Access-Publikationen in den Geistes- und Sozialwissenschaften mit dem Schwerpunkt Monografien, für die nächsten zwei Jahre finanziell gefördert.¹⁰

Abgerundet wurde der Block zu Open Access durch die Rückmeldungen aus allen vertretenen Häusern zur Aufstellung der eigenen Universität(sbibliothek) in diesem Bereich, die strukturiert anhand eines vorab verschickten Fragenkatalogs gesammelt wurden.

Wissenschaft und Wissenschaftsvermittlung

Der zweite Tag des Erfahrungsaustausches war der Wissenschaft gewidmet und wurde von André Welters (USB Köln) moderiert. Dabei standen am Vormittag Beiträge von Fachwissenschaftler/inne/n verschiedenster Disziplinen auf dem Programm, die anwendungsbezogen technische Forschung

7 OLH, s. <https://www.openlibhums.org/>, zuletzt geprüft am 27.04.2018.

8 S. <http://www.lingoa.eu/>, zuletzt geprüft am 27.04.2018.

9 Das Projekt Open APC veröffentlicht hierzu Daten aller beteiligten Institutionen, s. <https://treemaps.intact-project.org/>, zuletzt geprüft am 27.04.2018.

10 S. <http://www.uni-due.de/ogesomo>, zuletzt geprüft am 27.04.2018.

erläuterten, Wissenschaftstheorie und -geschichte nebst Wünschen an Bibliotheken beleuchteten oder die fachliche Neuausrichtung in einem anwendungsbezogenen Studiengang schilderten.

So stellte der Informatiker Dr. Rainer Stotzka vom Karlsruher Institute of Technology seine Forschungen zur „Automatischen Layoutanalyse mittelalterlicher Handschriften und frühneuzeitlicher Drucke“ vor. Diese Grundlagenforschung soll es in wenigen Jahren ermöglichen, dank eines automatisierten Verfahrens das Layout von in Masse digitalisierten Drucken und Handschriften zu beschreiben, also Schrift, Bild, Rand und sonstige Bestandteile eines Digitalisates zu unterscheiden. Zur Anwendung kommen dafür Röntgenstrahlen und mathematische Kurvenberechnungen aus den unterschiedlichen Farbsättigungsgraden zwischen Schrift, Illustration und Hintergrund. Durch grafische Darstellung der einzelnen Seiten kann es sogar möglich werden, interlineare Glossen oder unterschiedliche Handschriften und Schreibschulen zu identifizieren. Als Baustein eines größeren Forschungsverbundes dient diese Arbeit auch der Vorbereitung einer OCR-Software, die z.B. Frakturschrift zuverlässig erkennen soll. Dank dieses verständlichen und gut erläuterten Vortrags konnten Verbindungen zwischen zunächst völlig fachfremd erscheinenden Disziplinen geknüpft werden.

Dass es Bedarfe und Wünsche an Informationstechnologie zur Bilderkennung, Identifikation von Bild-Text-Digitalisaten und an Verfahren zur Recherche nach Bilddaten und Bildmaterial gibt, verdeutlichte der Vortrag von Prof. Dr. Jens Jäger von der Universität zu Köln. Er referierte über den „Iconic Turn – zum Visuellen in den Geschichtswissenschaften“ und beklagte, dass in den vergangenen großen Digitalisierungsbemühungen kein Augenmerk auf die Erfassung und Beschreibung von Bildern gelegt wurde, da die Geschichtswissenschaft als Profession lange eng am Text als einziger Überlieferungsform geblieben sei, und nun kaum systematische Recherchen zur Geschichtsschreibung mit Hilfe von Bildern oder zur Verbreitungs- und Rezeptionsgeschichte von bewusstseinsprägenden Bildern möglich seien.

Die „Wissenschaftsvermittlung am Beispiel von „Public History““ dagegen, so der Titel des Beitrags von Jun.-Prof. Dr. Christine Gundermann (ebenfalls von der Universität zu Köln), stellte den Versuch dar, im Fach Geschichte einen anwendungsbezogenen Studiengang zu implementieren. Dafür wurden US-amerikanische Konzepte und Inhalte nach Deutschland übertragen und eine wissenschaftliche Fundierung diskutiert. Die ersten Studierenden mit diesem Schwerpunkt gibt es derzeit an der Universität zu Köln, die Organisation liegt allerdings bei befristet beschäftigten Juniorprofessuren. Ein ähnlicher Studiengang ist mit der Fachjournalistenausbildung an der Universität Gießen zu finden. Für die Zusammenarbeit mit Bibliotheken von Bedeutung ist nach Ansicht der Referentin die auch in anderen Fächern zu beobachtende Ausdehnung der Forschungsgegenstände auf jegliche Veröffentlichungsform, insbesondere auch Populärschriften wie Comics, Illustrierte etc. und weitere Aspekte der Alltagskultur. Dadurch werde der Sammelauftrag der Bibliotheken immens erweitert und vor technische Herausforderungen gestellt, sollen doch auch Fan-Zeitschriften, Online-Rollenspiele oder Merchandising-Artikel untersucht werden.

Wissenschaftskommunikation und Services für die Wissenschaft

Am Nachmittag wurden in verschiedenen Praxisberichten aus dem Kreis der Teilnehmenden Möglichkeiten, Ideen und Wege zur Wissenschaftskommunikation beschrieben, die als Anregung für alle Fachreferentinnen und Fachreferenten dienen, in besseren Kontakt zu den von ihnen betreuten Forschenden und Lehrenden zu treten, sei es durch systematisch angebotene Dienstleistungen oder durch neue Kommunikationswege. Die Moderation dieses Blocks übernahm Dr. Rosemarie Kosche von der UB Duisburg-Essen.

Björn Gebert (HAAB Weimar) zeigte, ausgehend von seiner eigenen Redaktionstätigkeit im Mittelalter-Blog,¹¹ verschiedene „Portale, Kanäle und Werkzeuge digitaler Wissenschaftskommunikation“ auf, um „Fachcommunities und Fachreferat in den Geisteswissenschaften“ besser zu vernetzen, insbesondere für Forschende der Mediävistik. In Twitter beispielsweise seien die Schlagworte #icanhazpdf und #followerpower gut geeignet, um Texte auszutauschen oder die eigenen Themen zu verfolgen. In der Diskussion ergänzte Viola Voß aus Münster, über RSS-Feeds von ausgewählten Seiten, Blogs oder Twitter-Accounts relevanter Forschender sei es halb-automatisiert möglich, selber regelmäßige Neuigkeiten zu erfahren und an die eigenen Kunden weiterzuverbreiten. Auch professionelle Netzwerke wie Academia oder die neue Plattform Humanities Commons¹² fanden Erwähnung.

Thorsten Lemanski, Abteilungsleiter Digitale Dienste der ULB Düsseldorf und gleichzeitig Fachreferent für Philosophie, sieht die derzeitige Umstrukturierung seines Hauses als Chance, neue Angebote von Grund an aufzubauen und zu etablieren. So schilderte er seine Überlegungen zu „Services für die Wissenschaft im Bereich der Digitalisierung“: Im Rahmen des neuen Urheberrechtsgesetzes könnten Reproduktionsdienste angeboten und digitale Kopien oder Faksimiles im Digitalisierungszentrum hergestellt und für Wissenschaft und Forschung bereitgestellt werden. Auf diese Weise sieht er Potential für Kopienliefer- und Digitalisierungsdienste, die Forschungsapparate (analog zum Semesterapparat in der Lehre) unterstützen würden. Das Problem hierbei seien die noch ungeregelte Vergütung und die bekannte Tatsache, dass sich viele Wissenschaftlerinnen und Wissenschaftler nicht an urheberrechtliche Regeln hielten, deren Einhaltung die Bibliothek in einem solchen Service kontrollieren würde. Einige Anfragen zur Forschungsunterstützung aus verschiedenen Fächern im Kontext der Digital Humanities zeigten, dass hier Bedarf besteht und dass die bibliothekarische Kompetenz für Meta- und Normdaten sowie die Garantie der Langzeitarchivierung bekannt sind.

Nach einer letzten Kaffeepause übernahm Yvonne Brzoska (UB der RWTH Aachen) die Abschlussdiskussion zur Situation des geisteswissenschaftlichen Fachreferates. In einem einführenden Kurzreferat zum Strategie- und Umstrukturierungsprozess des wissenschaftlichen Dienstes in der UB Aachen stellte sie die derzeit überall geführte Diskussion ums Fachreferat in zwei Szenarien dar: (1) Die Konzentration aller Fachreferate auf wenige Mitarbeiterinnen und Mitarbeiter des wissenschaftlichen Dienstes, während die anderen sich auf wissenschaftsbezogene Dienstleistungen wie Open Access,

11 S. <http://mittelalter.hypotheses.org/>, zuletzt geprüft am 27.04.2018.

12 S. <https://hcommons.org/>, zuletzt geprüft am 27.04.2018.

Forschungsdatenmanagement oder Publikationsunterstützung spezialisieren (wie in Leipzig realisiert) und (2) die von vielen Häusern praktizierte Aufgabenverteilung mit Schwerpunkt-Spezialistinnen und Spezialisten, die dann zusätzlich ein oder mehrere Fächer im Fachreferat betreuen. Die Berichte aus den aktuellen Veränderungsprozessen ergaben Detailspekte wie die Organisation in Fächerclustern, die teilweise mit der Leitung von Institutsbibliotheken einhergehen, oder eine Gliederung in verschiedenen intensive Servicelevel der Fachreferatsarbeit je nach Gewicht des einzelnen Faches, und zeigten die manchmal schwierig zu verarbeitenden Veränderungen des Berufsbildes. Dabei ermöglichten der kollegiale Austausch und die Herkunft der Teilnehmenden aus verschiedensten Häusern den offenen Einblick in Chancen und Risiken dieser Veränderungen, die überall unter verschiedensten Rahmenbedingungen ablaufen.

Die Abschlussrunde mit Evaluation und Ausblick ergab große Zufriedenheit bei allen Anwesenden und den Wunsch nach Fortführung als zweitägige Veranstaltung für denselben Adressatenkreis, allerdings eher an einem anderen Ort als dem sehr attraktiven, aber für viele außerhalb Nordrhein-Westfalens recht teuren und abgelegenen Kardinal-Schulte-Haus. Neben verschiedenen Themen für das nächste Jahr nahm das Organisationskomitee damit den Auftrag der Fortführung mit und wird beizeiten ein neues Programm vorstellen.

Dorothee Graf, Universitätsbibliothek Duisburg-Essen

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S202-207>

Neue Konzepte für die Informations- und Literaturversorgung?

Eine Zusammenfassung der Fortbildung „Fachinformationsdienste: Angebote und Perspektiven der überregionalen Informationsversorgung“

Am 12. Oktober 2017 veranstaltete der VDB-Landesverband Bayern in der Universitätsbibliothek Erlangen-Nürnberg in Erlangen eine ganztägige Fortbildungsveranstaltung zum Thema Fachinformationsdienste für die Wissenschaft.



Die Überraschung war auf Seiten der Wissenschaftlichen Bibliotheken zunächst groß, als die DFG nach vorangegangener Evaluation das seit mehr als sechzig Jahren bestehende System der überregionalen Literaturversorgung in Form der Sondersammelgebiete in drei Etappen, beginnend mit dem Jahr 2014, beendete und durch das Konzept der Fachinformationsdienste für die Wissenschaft ersetzte. Ging diese Neukonzeption doch zugleich mit einem wesentlichen Paradigmenwechsel einher: So sollten die projektorientierten und in enger Kooperation mit der jeweiligen Fachcommunity ausgerichteten Fachinformationsdienste das einheitliche und nach verbindlichen Kriterien aufgebaute System der überregionalen Literaturversorgung ersetzen. Zudem sollte der Fokus zukünftig auf der Entwicklung elektronischer Dienstleistungen liegen, die auch in der Erwerbung den elektronischen Medien Vorrang vor den konventionellen Medien einräumt und sich nunmehr auf den wissenschaftlichen Spitzenbedarf eines Faches konzentriert.

Betraf dieser Umstieg im ersten Jahr des neuen Förderprogramms zunächst einige Geistes- sowie die Sozialwissenschaften, so waren im zweiten Jahr die Naturwissenschaften und im dritten eine große Gruppe geisteswissenschaftlicher und ehemaliger geographisch-regionaler Sondersammelgebiete an der Reihe. Allein mit einer Aktualisierung oder Fortführung bisheriger Erwerbungsprofile war es nicht getan, denn es verlangte nach neuen Konzepten und inhaltlichen Neuausrichtungen, die im engen Austausch mit der jeweiligen Fachwissenschaft erarbeitet wurden. So finden sich vielfach auch neue Partner und Allianzen unter dem Dach eines Fachinformationsdienstes wieder, der jetzt häufig aus der Zusammenarbeit von Bibliothek und einer auf IT und Datenverarbeitung ausgerichteten Informationseinrichtung besteht.

Während die Antragssteller in den ersten beiden Transformationsphasen mangels erfolgreicher „Best-practice-Anträge“ in vielen Fällen erst im zweiten Versuch erfolgreich waren, lag die Förderquote in der dritten Runde durchweg höher. Dennoch gibt es kein allgemeines Erfolgsrezept, da die Begutachtungsprozesse höchst individuell waren und sich daraus disziplinübergreifend keine Empfehlungen für einen erfolgreichen Antrag ableiten ließen.

Diese Bandbreite und Heterogenität der Fachinformationsdienste demonstrierten auch die sieben Beiträge dieses Fortbildungstages. Den Anfang machte Dr. Gregor Horstkemper (BSB München) mit seinem Vortrag „Zwischen Literaturversorgung und Dienstentwicklung: Akzentsetzungen im FID Geschichtswissenschaft.“ Dieser, von der Bayerischen Staatsbibliothek und dem Deutschen Museum gemeinsam getragene Fachinformationsdienst, erörterte die Chancen der FIDs, die in einem gemeinsam mit der Fachcommunity erarbeiteten Informations- und Literaturangebots liegen. Allerdings wurden auch die damit verbundenen Schwierigkeiten thematisiert, die im Falle der Geschichtswissenschaft in der Größe und Heterogenität der Fachdisziplin sowie in deren Interesse an einem vorsorgenden Bestandsaufbau, gerne auch mit gedruckter Literatur, lagen. Auch wurde ein im Rahmen des FID nur auf Wissenschaftler/-innen beschränkter Zugang zu elektronischen Medien als unzureichend erachtet. Alles Aspekte, die der offiziellen DFG-Policy eigentlich widersprechen. So versucht der FID Geschichtswissenschaft, die konkreten Bedürfnisse der Wissenschaft durch folgende Arbeitsschwerpunkte umzusetzen:

1. Ein weiterhin vorsorgender Bestandsaufbau, der sich bei Zeitschriften allerdings nunmehr auf unikale und singuläre Bestände beschränkt und für Monographien im Rahmen eines inhaltlichen Erwerbungsprofils erfolgt.
2. Der Abschluss zielgruppenspezifischer FID-Lizenzen des wissenschaftlichen Spitzenbedarfs mit Nutzungskontingenten für Datenbanken und E-Books, der den Bedürfnissen der Historiker nach Interdisziplinarität Rechnung trägt und Zugriff auch für Studierende bietet.
3. Die Weiterentwicklung des Recherche- und Informationsportals *Historicum.net* mit fachspezifischen Einstiegswegen.
4. Die Etablierung einer deutschen historischen Bibliographie, die gemeinsam mit weiteren Projektpartnern realisiert werden soll sowie
5. Serviceangebote rund um die Digitalisierung von Quellenbeständen und das elektronische Publizieren.

Thema des zweiten Vortrags von Dr. Martin Faßnacht (UB Tübingen) war „Der neue FID Theologie. Schwerpunkte – Organisation – Neuheiten“. Kumulationspunkt des Fachinformationsdienstes ist der Index Theologicus *ixtheo.de*, der in Kooperation mit weiteren Bibliotheken, Verlagen und anderen Bibliographien zum zentralen Nachweisinstrument für die Erwerbung und Erschließung theologischer Literatur ausgebaut wird. In diesem Sinne wurde *ixtheo.de* um zusätzliche Suchfeatures wie eine Bibelstellensuche, die thematische Suche in verschiedenen Sprachen sowie einen Alerting Dienst für neu erschienene Aufsätze erweitert. Darüber hinaus besitzt der FID auch weiterhin eine wichtige Erwerbungs-komponente, in dem die UB Tübingen sowohl auf Nutzerwunsch als auch im Sinne der bibliographischen Vollständigkeit des Index Theologicus monographische Literatur im größeren Umfang erwirbt, so dass in diesem Bereich der vorherige Erwerbungs-auftrag des Sondersammelgebiets Theologie in die Gegenwart überführt wurde, nun aber ergänzt um die Komponente der elektronischen Medien. Dies schließt auch Nationallizenzen für kleine, fachliche definierte Nutzergruppen ein. Ein zukünftiger Schwerpunkt liegt zudem im Ausbau der elektronischen Services. Dies erfolgt durch eine retrospektive Digitalisierung historischer Werke und den Aufbau von Publikationsdienstleistungen im Rahmen der Zweitveröffentlichung von Aufsätzen in Open Access.

Gänzlich andere Herausforderungen stellen sich dem neuen regionalen FID Osteuropa. So verwies Dr. Gudrun Wirtz (BSB München) in ihrem Beitrag „Neuer Wein in neuen Schläuchen? – Die Dienste der BSB für die Ost-, Ostmittel- und Südosteuropaforschung und ihre Förderung durch die DFG“ auf die konzeptionellen Anforderungen, die sich aus der Integration dreier ehemaliger Sondersammelgebiete in einen Fachinformationsdienst ergeben. Dies betrifft einerseits die komplizierte Verquickung aus regionaler und fachlicher Zuordnung als auch die spezielle Situation der Literatur- und Informationsversorgung in Ost-, Südost- und Ostmitteleuropa, die weiterhin einen vorsorgenden Bestandsaufbau mit primär gedruckter Literatur erfordert. Auf Grund der geringen Angebotsvielfalt sind elektronische Angebote nur begrenzt verfügbar, so dass in diesem Segment eine E-Only-Policy nicht realisiert werden kann. Andererseits erfordert die politische Situation in einigen Staaten Osteuropas die Archivierung relevanter Open-Access-Journals durch den FID, um diese vor Zensur zu bewahren. In diesem Zusammenhang kommt auch der Virtuellen Fachbibliothek „Vifa Ost“ eine besondere Rolle zu, deren Angebote zudem um die Bereiche Elektronisches Publizieren und Retrodigitalisierung ausgebaut werden. Der Ausrichtung auf die neuen Förderrichtlinien der DFG trägt der FID durch eine Profilschärfung im Bestandsaufbau Rechnung. So konnten zwar die Erwerbungs-ausgaben durch Absprachen mit verschiedenen Spezialbibliotheken reduziert werden, dennoch erfordert die spezielle Situation des osteuropäischen Buchmarktes und der damit verbundenen Erschließungsprobleme einen hohen Eigenanteil der Bibliothek an Erwerbungs- und vor allem Personal-mitteln, die eine Betreuung dieses FID finanziell höchst unattraktiv machen.

Einen anderen Ansatz verfolgt hingegen der FID Geowissenschaften der festen Erde (FID Geo), der gemeinsam von der SUB Göttingen und dem Deutschen GeoForschungsZentrum (GFZ) in Potsdam betrieben wird. So erläuterte Dr. Norbert Pfurr (SUB Göttingen), dass man sich in diesem Gemeinschaftsprojekt ganz bewusst gegen eine kostenpflichtige Erwerbung von Fachliteratur im FID-Kontext entschieden hat und sich stattdessen in Absprache mit den einschlägigen wissenschaftlichen Fachgesellschaften auf einen Ausbau der E-Services und den Nachweis von Open-Access-Publikationen konzentriert. Dazu zählen sowohl der elektronische Zugriff auf graue sowie open access verfügbare

Literatur im FID-eigenen Repository Geo-Leo als auch die Bereitstellung und Archivierung von Forschungsdaten durch das GFZ Potsdam. Weitere Angebote des FID sind die Digitalisierung on Demand für gemeinfreie Werke, institutionelle Serien und Karten sowie die Beratung von Autor/-innen und Wissenschaftler/-innen in allen Fragen rund um das elektronische Publizieren und die Archivierung von Forschungsdaten. Während sich die Beratungsdienstleitungen schon nach kurzer Zeit einer hohen Nachfrage erfreuen, gestaltet sich die Digitalisierung institutioneller Serien und Karten infolge der vorab notwendigen Rechteeinholung schwierig. Wie überhaupt die Kommunikation mit einer aus mehr als 20 Fachgesellschaften bestehenden Fachcommunity besondere Herausforderungen stellt und derzeit im Wesentlichen über eine Mailingliste und Informationen in den Geowissenschaftlichen Mitteilungen (GMIT) und auf Fachtagungen erfolgt.

Mit seinem Vortrag „FID Biodiversitätsforschung – ein FID mit Fokussierung auf bedarfsorientierten Text-Mining“ präsentierte Dr. Gerwin Kasperek (UB Frankfurt a.M.) einen zweiten Fachinformationsdienst aus den Naturwissenschaften. So hatte man nach vorangegangenen Experteninterviews innerhalb der Fachcommunity der Biologie vor allem in der Biodiversitätsforschung einen hohen Informationsbedarf ermittelt und die vier verschiedenen Module Text-Mining, Open Access, Digitalisierung und Literaturerwerbung als die wesentlichen bedarfsorientierten Komponenten des FID identifiziert. Im Mittelpunkt des Text-Mining und der Digitalisierung steht ein Pilotprojekt zur Überführung historischer Textinformationen zum Vorkommen von Tier- und Pflanzenarten des deutschsprachigen Raums in eine Datenbank. Dazu werden die Texte zunächst digitalisiert, deren Daten mittels Software retrievelfähig eingespielt und der Forschung zur Nachnutzung zur Verfügung gestellt. Ein weiterer Schwerpunkt ist die Überführung relevanter, von Fachgesellschaften herausgegebener Zeitschriften in Open Access und deren Präsentation auf einer OJS-Plattform. Anders als in den Geowissenschaften spielt jedoch die Erwerbung im FID Biodiversitätsforschung nach wie vor eine wichtige Rolle: Im Abschluss relevanter FID-Lizenzen, in der Bereitstellung von Artikeln aus mehr als 550 Fachzeitschriften sowie im Erwerb monographischer Literatur. Der Austausch mit der Wissenschaft erfolgt mittels Beirat, Workshops und Rundmails an verschiedene Empfängerkreise innerhalb der Fachcommunity. Ob sich der Fachinformationsdienst allerdings in diesem Kontext als „Marke“ etablieren kann, sieht Dr. Kasperek in Anbetracht der internationalen Ausrichtung der Biodiversitätsforschung und der Nischenausrichtung auf den „Spitzenbedarf“ kritisch.

Einen Fachinformationsdienst der ersten Stunde präsentierte Dr. Sebastian Stoppe unter dem Titel „Was sich hinter *adlr.link* verbirgt: Informationsdienstleistungen und Literaturversorgung für die Kommunikations-, Medien- und Filmwissenschaft.“ Mit ca. 1,3 Mio. Datensätzen ist *adlr.link* das zentrale Rechercheportal des FID Kommunikations- und Medienwissenschaften und bietet neben Büchern und Zeitschriften auch den Zugriff auf E-Books und zahlreiche audio-visuelle Daten in Form von Bild- und Musikdaten. Seinen Nutzer/-innen bietet *adlr.link* nach Anlegen eines kostenlosen Nutzerkontos zahlreiche personalisierte Such- und Arbeitsfunktionalitäten, darunter auch die Services eines kostenlosen Kopienversands von Artikeln aus der Datenbank und einer bedarfsorientierten Erwerbung von Monographien und deren Lieferung an den Endnutzer, sofern dieser Titel nicht an der Heimatbibliothek des Bestellers / der Bestellerin zur Verfügung steht. Auch werden Verlagsneuerscheinungen auf Basis eines Approval Plans in die Datenbank integriert und können zur Anschaffung und späteren Direktlieferung vorgeschlagen werden. Darüber hinaus hat der Fachinformationsdienst

im Bereich der elektronischen Medien mit dem Abschluss nutzergruppenspezifischer FID-Lizenzen sein Angebot ausgeweitet und auf die Bedürfnisse seiner Fachcommunity ausgereicht.

Zum Abschluss der Fortbildung erläuterte Dr. Rainer Plappert (UB Erlangen-Nürnberg) mit seinem Vortrag „Zwischen Spitzenbedarf und E-Only-Policy. Der FID Erziehungswissenschaft und Bildungsforschung und die Bedürfnisse der Fachcommunity.“ Spezifika dieses FID. So nimmt der FID Erziehungswissenschaft und Bildungsforschung eine Sonderstellung ein, da er von fünf Projektpartnern in vier Institutionen getragen wird. Neben dem hohen administrativen Aufwand stellt auch die Kommunikation mit der zahlenmäßig großen und sehr heterogenen Fachcommunity eine besondere Herausforderung dar. Allerdings verfügt der Fachinformationsdienst mit dem Fachportal Pädagogik über ein bereits seit vielen Jahren etabliertes und in der Fachwelt gut eingeführtes Informations- und Rechercheportal, über welches die zentralen Dienste der bedarfsorientierten Erwerbung und Digitalisierung on Demand bereitgestellt werden. Zugleich wurde im Zuge der Einführung des Fachinformationsdienstes die Recherche nach erziehungswissenschaftlicher Literatur durch die Einbeziehung weiterer Datenlieferanten, wie der Library of Congress und des Bibliothekslieferanten Casalini, verbessert. Diese Datenlieferungen bilden die Grundlage einer bedarfsorientierten Monographienerwerbung der in Deutschland überregional nicht vorhandenen ausländischen Fachliteratur. Darüber hinaus erwirbt der FID die unikale und in den vergangenen Jahren auch tatsächlich nachgefragte Zeitschriftenliteratur und versucht, diese in Form von Nationallizenzen überregional zur Verfügung zu stellen. Die Praxis zeigt jedoch, dass sich das von der DFG propagierte Primat einer E-Only-Policy in einer noch deutlich durch gedruckte Medien bestimmten Angebotsstruktur der Erziehungswissenschaften nur in Ansätzen realisieren lässt.

Die eintägige Fortbildung vereinte eine Reihe ganz unterschiedlicher Fachinformationsdienste mit ihren unterschiedlichen Charakteristika. Dabei wurde deutlich, dass sich das jeweilige Angebotsspektrum als das Ergebnis eines mehr oder minder intensiven Dialogs mit der jeweiligen Fachcommunity präsentiert. Die Bandbreite der Angebote, unter denen die Literaturversorgung nur noch eine von vielen Optionen ist, reicht von Open-Access-Aktivitäten über sehr differenzierte Erwerbungsprofile, erweiterte Suchportale, Data-Mining, Social-Media-Präsenz bis hin zum Forschungsdatenmanagement. Vielen Fachinformationen gemeinsam sind dabei die umfangreichen Serviceangebote rund um das elektronische Publizieren.

In der abschließenden Diskussion war allerdings auch ein Unbehagen spürbar, welches für manche Teilnehmenden mit der Abkehr von einer traditionell langfristigen flächendeckenden Infrastrukturförderung hin zu einer punktuellen und projektorientierten Förderung verbunden ist. Die aktuell fehlende Verstetigung dieser grundlegenden Infrastrukturaufgabe wirft Fragen auf, die auch über die Zukunft dieses FID-Systems entscheiden. Wird es den Fachinformationsdiensten gelingen, sich dauerhaft als Partner der Wissenschaft zu etablieren, werden sie akzeptiert und vor allem, werden ihre Dienstleistungen auch in Anspruch genommen? Gerade in Zeiten, in denen die Kosten-Nutzen-Relation eine immer größere Bedeutung erhält, stellt sich die Frage, in wieweit ein auf den Spitzenbedarf ausgerichteter Dienst sich durch die umfangreiche Nutzung seiner Angebote legitimieren muss. Oder andersherum: Wie wenig Nutzung wird toleriert? Welchen Mehrwert bieten also die von ihnen angebotenen Dienstleistungen, und ist eine Verengung auf den aktuellen Spitzenbedarf der

Wissenschaft im Sinne des Aufbaus nachhaltiger Informationsinfrastrukturen wirklich zielführend? Schließlich könnte sich auch für die Träger dieses Systems die Frage stellen, welche Aufgaben den Bibliotheken bei einer Abkehr von einer vornehmlich bestandsorientierten Literaturbeschaffung zukünftig noch in diesem FID-System bleiben. Dies waren nur einige der Fragen, mit denen die Fortbildung in einer regen Diskussion ihr Ende nahm.

Rainer Plappert, UB Erlangen-Nürnberg

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S208-213>

Berichte und Mitteilungen

Research Information Systems – fit for the future?

A report on the situation and plans of the University of Sheffield Library

1. Introduction

This article discusses the University of Sheffield's rationale for joining, and the experience of participating in, the Development Partner Programme for Ex Libris' Esploro product. It should be noted that the article expresses the opinions of the University of Sheffield Library staff writing the article, and not the Ex Libris company.

The University of Sheffield is located in South Yorkshire in the North of England. It was founded in 1905 from a merger of local technical and medical colleges and has a prominent role in the region and city as a civic university. Sheffield is a research-intensive institution with a background in research-led teaching and learning. There is a strong relationship with local industry and the regional economy, most notably manifest in the Advanced Manufacturing Research Centre located just outside Sheffield.

The University has 25.000 full-time students, almost 3.000 undertaking part-time study. Of these, over 2.700 are studying for a degree by research (MPhil or PhD). There are approximately 1.400 academic staff across five subject-based faculties: Arts & Humanities, Social Sciences, Science, Engineering, and Medicine, Dentistry & Health. Additionally, the International Faculty is based in Thessaloniki, Greece. The University has a strong track record in research and was joint 14th on grade average in the most recent UK Research Excellence Framework (assessing the quality of research outputs).¹ In the QS World University Rankings, Sheffield ranks 82nd in the world and 13th in the UK.²

2. Current Research Environment

The University teaches and researches across a wide range of disciplines in addition to significant interdisciplinary research and collaborations from researcher level through the regional UK N8 group to strategic international partnerships. All this research activity leads to diverse requirements in terms of research information management (RIM) and research data management (RDM).

A significant proportion of research is funded through UK Research Councils (RCUK, now newly amalgamated under UKRI, UK Research and Innovation) or large charity funders such as the Wellcome Trust. This complex funding landscape leads to specific, and occasionally diverse, requirements for that impact on RIM and RDM. Policy and mandates underpinning grant applications, data management plans, persistent researcher identifiers (one variant being ORCID), institutional identifiers, and

1 "Research Excellence Framework 2014: Overall ranking of institutions," Times Higher Education, accessed 20 April 2018, <https://www.timeshighereducation.co.uk/sites/default/files/Attachments/2014/12/17/k/a/s/over-14-01.pdf>.

2 "QS World University Rankings," QS Top Universities, accessed 20 April 2018, <https://www.topuniversities.com/university-rankings/world-university-rankings/2018>.

open access publishing requirements can vary to a significant degree between funders. In addition to funder requirements, universities and researchers must comply with emergent government policy around open science and open access to research outputs.³ Hitherto, this has not necessarily been consistent with individual funder policy, the most prominent example of this being the divergent routes to open access research articles taken by RCUK and HEFCE (Higher Education Funding Council for England), though with RCUK transitioning to become UKRI, it may be converging.⁴ Open access is generally seen as desirable but the routes taken to achieve this goal are not necessarily the same. Over and above this, the University has to take into consideration the international landscape, most notably the EU, where policy although consistent in the general aims can be subtly at variance with regard to the detail around open access and research data.

Another key part of the UK higher education landscape is the Research Excellence Framework (REF)⁵ which is an exercise held approximately every six to seven years which monitors the quality of research outputs from institutions and determines apportionment of government research funding for institutions. It is imperative that institutions with a research focus do as well as possible in this exercise. In order for research outputs to be considered for the REF exercise they must be openly accessible to the public, either by paying for 'gold' open access, or by making the author-accepted manuscript available in a repository within a certain timescale, i.e. 'green' open access. All of this, along with necessary exceptions, requires managing in a systematic way.

3. Existing Systems Infrastructure

The research environment and diverse research outputs created (such as pre-print articles, published articles, theses, conference papers, monographs and datasets) has led the University to implement a range of systems to help manage these objects and provide deposit workflows and end-user discovery. Outputs have multiple facets that require coordinated management including discovery, financial management, funder compliance, links to grants and attributions to both researchers and institutions. In the past the systems infrastructure, processes and workflows used to manage outputs have developed in response to policy and the external environment. Inevitably, this has resulted in integration challenges, a fragmented overall architecture and in some cases gaps where key functionality is now mission critical.

3 "Government to open up publicly funded research," UK Government, published 16 July 2012, <https://www.gov.uk/government/news/government-to-open-up-publicly-funded-research>

4 "Open access," UK Research and Innovation, accessed 20 April 2018, <https://www.ukri.org/funding/information-for-award-holders/open-access/>.

5 "REF2021," HEFCE, Research England et al., accessed 20 April 2018, <http://www.ref.ac.uk/>.

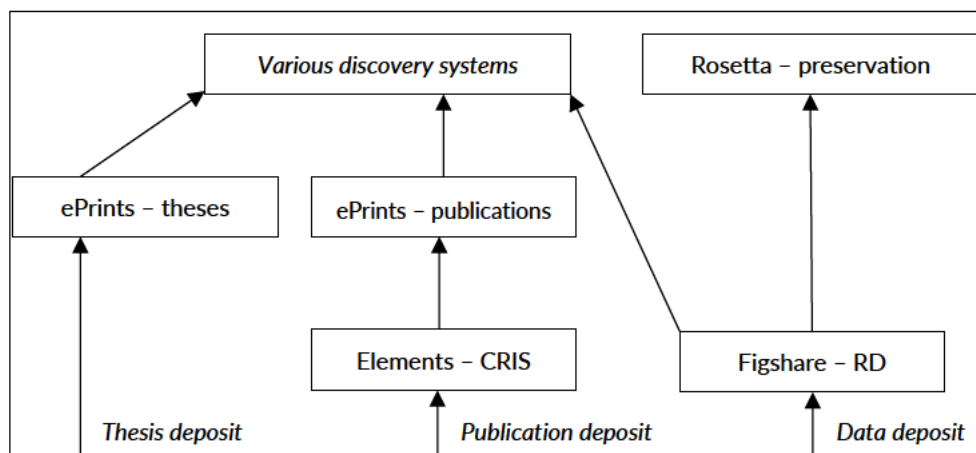


Figure: Simplified diagram of data/metadata flows in University of Sheffield systems discussed in this article

3.1. Institutional Repository (IR)

The University Library manages the institutional repository (IR) which in the case of Sheffield is an ePrints instance jointly run as a shared service with the regional White Rose Consortium.⁶ The IR was originally an output of the Jisc SHERPA Repositories Support programme and since its inception has now become a business critical part of the research systems environment. Deposits in Symplectic Elements are harvested into the repository, where Library staff check open access compliance and update metadata accordingly. The articles can then be published in the repository, fulfilling a range of public open access mandates and meeting compliance targets. Thereafter outputs are then harvested by other discovery tools including the University Library's Primo resource discovery solution and repository aggregator services.

3.2. Electronic Theses Repository

Institutional eTheses are also held in a shared service repository with White Rose. Deposit of theses (post-examination and pre-degree award) is handled by a different instance of ePrints (White Rose eTheses Online - WREO). This has a customised deposit workflow to handle specific eTheses requirements, such as supervisor details and embargoes at multiple content levels.

3.3. CRIS

For deposit of open access articles (in particular the author's accepted manuscript), the University uses Symplectic Elements.⁷ Deposit can be a manual process, though Elements has integrations into Web of Knowledge, Scopus, arXiv and with ORCID, in order to harvest metadata and allow researchers to 'claim' publications removing the necessity to re-key metadata. The University's research office is the lead stakeholder for this system as Elements is used to support managing the REF exercise.

6 "Eprints," Eprints, accessed 20 April 2018, <http://www.eprints.org/uk/>.

7 "Elements," Symplectic, accessed 20 April 2018, <https://symplectic.co.uk/products/elements/>.

3.4. Research Data Management

Some funders now require management and publishing of the data underlying the research that is undertaken.⁸ The first step in managing research data is the production of a Data Management Plan (DMP). The Library encourages researchers to use tools such as DMPonline⁹ to create DMPs, but at present there is no integration with any local systems.

Once the dataset is ready for publishing, researchers have a number of options. There are many subject-based data repositories available (for example, the UK Data Service, or the Archaeology Data Service) that the Library encourages researchers to use if they are suitable. These meet funder requirements, but make it very difficult for the University to gain a good picture of the range of its research outputs or to showcase them. In addition, there are many subjects for which there is no suitable external data repository. In response to this fragmented picture, to meet compliance mandates and to underpin showcasing of data the University acquired figshare for institutions.¹⁰ Researchers can deposit datasets into figshare, or create metadata-only records if the dataset is held in an external repository or if the data is too large or sensitive for holding in an online repository. Currently there is no way to harvest external data repository records into figshare, meaning that researchers must duplicate their efforts.

In order to improve the showcasing of the University's research outputs, the Library has collaborated with the IT department to create a custom front-end to figshare that is called ORDA. This displays research data outputs in an attractive way and is styled to match University branding. ORDA is a thin web front-end built using figshare APIs. It has also been possible to add additional services in ORDA. For example, if a researcher has created a visualisation of their dataset (using plot.ly for example), they can add a metadata file to their dataset which indicates the location of the visualisation and ORDA can then display this alongside the dataset.

3.5. Preservation of Research Data Outputs

The University Library is also concerned about preservation of digital assets including research outputs. The University's Digital Preservation function works within the Library, and the University has procured the Rosetta digital preservation system from Ex Libris.¹¹ This already preserves much material from the Library's Special Collections and archives, but work is now underway to link Rosetta with figshare and with ePrints to preserve research outputs (articles, theses and datasets). A number of challenges remain, especially with large datasets, but significant progress has already been made.

8 "Research Funder Policy Summaries," University of Sheffield Library, accessed 20 April 2018, <https://www.sheffield.ac.uk/library/rdm/funders>.

9 "DMPonline," Digital Curation Center, accessed 20 April 2018, <https://dmponline.dcc.ac.uk/>.

10 "figshare for institutions," figshare, accessed 20 April 2018, <https://figshare.com/services/institutions>.

11 "Rosetta," Ex Libris, accessed 20 April 2018, <http://www.exlibrisgroup.com/products/rosetta-digital-asset-management-and-preservation/>.

4. The Challenge

With this complex patchwork of systems come a number of problems. The connections between systems and the data flows have become very complex and difficult to manage. There are still some gaps in functionality, and areas where there is no way of managing all aspects of a process in one place. The channels of information about research, for example grants, opportunities, bibliometric data, are also limited.

For example, staff in the Library who support open access work have no way of fully managing the customer process, from the point that a researcher may contact them to the final publication of the open access copy. Different systems are involved in different parts of the work and the integrations are not good enough to allow an enquiry to be followed without staff doing significant work.

A particular difficulty with open access is the management of Article Processing Charges (APCs). Many research articles are published open access by the payment of a publication charge to the publisher (this is known as 'gold' open access). This can happen in hybrid journals where gold open access articles sit alongside articles that reside behind a subscription paywall. Alternatively, gold articles can be within journals whose only commercial model are the charges received for publication of gold open access articles. In recent years, the University Library has managed a block grant to support payment of these charges, but there are also other sources of funding. An important question for institutions is what this means for the total cost of ownership of the subscriptions.¹² Is the Library paying twice – both to publish articles and to access them? Subscriptions and the payments for access are managed via the Library's management system (Ex Libris Alma). However, there is no system which manages APCs and sources of APC funding. This is currently managed via spreadsheets. This makes analytics about the full cost of subscriptions extremely difficult. The fragmented nature of this workflow also builds complexity into the system which in turn results in higher overheads than those that would typically result from a unified and integrated system.

Additionally, links to external sources of data are not always available for some systems. In particular, it is not currently possible to harvest records about datasets published in repositories outside the University. There is a vast amount of information available that could be used to underpin the University's research enterprise, but at present there is no structured way to pull this into existing systems and to make it usable and meaningful in a resilient and sustainable way to key internal stakeholders – most importantly researchers.

Gaps in functionality also exist around linking outputs to grants, sources of funding, potential collaborations and identifying research strengths and weaknesses elsewhere. There is no one system where researchers can go to establish bibliometric information. They are faced with a variety of

¹² Cf. Stephen Pinfield, Jennifer Salter and Peter A. Bath, "The 'Total Cost of Publication' in a Hybrid Open-Access Environment: Institutional Approaches to Funding Journal Article-Processing Charges in Combination with Subscriptions," *Journal of the Association for Information Science and Technology*, 67, no. 7 (July 2016): 1751–1766, <https://doi.org/10.1002/asi.23446>.

metrics produced by different sources which they have to consult individually – this, as with all overly complex spaces, is time consuming.

There are difficulties in finding open access copies of articles, even when these exist. Tools, such as OA Button, aim to help users find copies of articles that they can access. However, the challenges ahead were highlighted when one of the authors participated in a recent test of several common tools. This was used on a test repository built for the purpose containing records from a wide array of sources, 100 % of which had an open access copy available somewhere. Results were not encouraging and suggest that there are problems to be solved. This is not just the traditional library issue of enabling users to find content, but ensuring that the institution's outputs can be found by users around the world.

5. Esploro

For all the reasons outlined above, there is a desire to bring together a number of different systems – exactly which ones depends upon what is technically and organizationally possible. The hope is to increase efficiency, and therefore free up resources for new initiatives (the data visualisation service being a recent example). It should make essentials, such as links between different types of research outputs (for example, a dataset and the publications that cite it) much easier than is currently possible. In addition, it should be possible to take workflows, e.g. for APCs, which are not managed within current systems, and integrate them properly.

The University is already a customer of Ex Libris, taking products such as the Alma Library Management System, the Primo Discovery System and the Rosetta Digital Preservation System. Therefore, there was interest when Ex Libris announced that they would be developing a research platform, now known as Esploro.

Ex Libris aim to develop a system that brings the workflows for different research objects (such as articles, datasets and theses) together. Whilst each type needs to be handled differently, bringing them together enables links between them to be easier to manage and more explicit. Ex Libris also hope to leverage the content that they have access to via ProQuest in order to make Esploro able to do as much automated discovery of research outputs as possible, reducing the effort required to deposit within the system. In particular, discovery of externally deposited datasets is a facility the University's systems do not currently offer.

Esploro would also be built with open access compliance in mind, and Development Partners from UK were recruited to bring the particular experience of UK compliance and feed that into the development. As a library systems company, Ex Libris has experience in the area of discovery, both with institutional-based discovery systems and harvesting of institutional records for external discovery systems. The discovery of the research outputs placed into Esploro will be a key part of the system.

Providing report and analytics from research systems is an essential requirement, both for reporting compliance to external institutions and for measuring performance internally. Ex Libris wanted to

work with their Development Partners to understand what reports and metrics were required in such a system and pressures on researchers are high.

As Esploro promises to meet a number of challenges the University faces with some of its research systems, and as Development Partners have an opportunity to influence development requirements to ensure that the system is fit for purpose, the University decided to join the Development Partner programme.

6. Development Partners and the Development Process

The Development Partners for Esploro are the Universities of Sheffield and Lancaster in the UK, and the Universities of Iowa, Miami and Oklahoma in the USA. The programme started in mid-2017 with Ex Libris presenting their vision for the product and doing an initial requirements gathering exercise. This was followed up by a number of webinars in different 'tracks' (i.e. topics) to discuss system requirements in more detail with staff who specialise in different areas of the research publication support.

Late in 2017, Ex Libris presented their proposal for what requirements should be included within the initial version of the product, i.e. the key elements that would make a system usable in at least one area. Development of this is underway and Partners expect further webinars and email conversations with Ex Libris analysts to clarify earlier comments and suggestions. Key priorities for the initial release include the deposit workflows for publications and datasets, automatic capturing (and claiming) of metadata records for publications, APC management, institutional repository functionality and publication to discovery services. As at the time of writing, Development Partners are expecting to be able to start testing releases of code from June 2018 onwards.

Ex Libris hope to have a system that Partners can implement early in 2019. Meanwhile, Ex Libris will be launching an Early Adopter programme, to expand the number of institutions that will contribute ideas and will help with the continuing development of the product.

As part of its contribution to the Development Partner Programme, the University of Sheffield has brought together a project team with expertise from the Library, the IT department and the research office, and experience of handling different types of research outputs. Some staff have experience of the deposit workflows and managing ingest of material, and others with discovery. Some staff have attended face-to-face meetings that have been held in the UK with Sheffield, Lancaster and Ex Libris, while others have only attended webinars for their specialist areas (the webinars feature participants from all Partners, plus Ex Libris).

Unlike many projects, the time required for this project is very uneven. During periods of requirement gathering, there can be quite a few meetings and webinars, and at other times there can be very little to do. Also, there have been very few meetings of the local project team: Most of the work so far has been to share Sheffield expertise with Ex Libris. There have just been a few local meetings to make sure everyone is aware of how the overall development is progressing.

There is a benefit just being involved as a Development Partner. There is the opportunity to reflect upon local practices, especially where workflows cross departmental boundaries. In a few cases, this has led to further meetings being set up outside the Esploro project to improve current procedures, without even requiring a new system. It is also possible to gain ideas from fellow Partners as they often have different ways of working, and Sheffield can look to adapt current procedures.

7. The Future

In the longer term, it is hoped that Sheffield's contribution will help Ex Libris to produce a research object management system that will meet requirements better than the current range of systems. As Esploro develops, Sheffield will be able to decide if it wishes to procure the system. This would probably involve replacing just one or two systems initially (integrating with the rest) with further expansion possible over time. Symplectic Elements plays a key role in the way the University manages its submission to the REF exercise, so this system could not possibly be replaced during the current cycle (due to end 2021), however, Esploro could integrate with it as ePrints currently does to provide a repository. Research data management and APC handling are other possible features to implement should the University decide to do so. It is hoped that the University of Sheffield's management of research outputs can then become increasingly efficient and that it will be possible to continue to deliver innovative services in this area.

Tracey Clarke, Andy Bussey, The University of Sheffield Library

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S214-221>

Bericht aus der 73. Sitzung der Arbeitsgemeinschaft der Verbundsysteme am 15. und 16. November 2017 in München

Auf Einladung des Bibliotheksverbunds Bayern (BVB) traf sich die Arbeitsgemeinschaft der Verbundsysteme zu ihrer 73. Sitzung am 15. und 16. November 2017 in München. Der folgende Bericht hat den Stand November 2017.

Alma

Die cloudbasierte Bibliothekssoftware der Firma Ex Libris, mit der alle Abläufe – von der Erwerbung über die Katalogisierung bis hin zur Ausleihverbuchung – verwaltet werden können, wird von der Österreichischen Bibliothekenverbund und Service GmbH (OBVSG) schrittweise implementiert.

Die sechs Bibliotheken, die zur 1. Kohorte gehören, sind am 21. August 2017 mit Alma in Betrieb gegangen. Zurzeit erfolgt die Implementierungsphase für die Kohorte 2 (sieben Bibliotheken und OBVSG). Überführt wurden auch die laufenden zentralen Dienste im OBV. Die Berichtsapplikation WUW der Wirtschaftsuniversität Wien und DISS-DB (Opus) werden mit dem Umstieg auf Alma eingestellt; der Dienst SFX läuft mit Aleph aus. Durch ein Neuladen/Neuindexieren konnten die Vorteile der deutschsprachigen Aleph-GND nach Alma erfolgreich überführt werden.

Mit der Migration der Verbunddatenbank soll das Projekt im März 2018 abgeschlossen sein. Die konkreten Überlegungen und Ressourcenplanungen für den Umstieg weiterer Einrichtungen wurden „eingetaktet“ und weitere Zeitfenster sind bis August 2019 reserviert. Bis Ende 2020/2021 soll das Projekt komplett abgeschlossen sein.

Im Zeitplan sind die beiden im April vom BVB gestarteten Projekte zur Erprobung von Alma und Bibliothekssystemen der nächsten Generation auf Open-Source-Basis. Für das Alma-Erprobungsprojekt hat im Oktober 2017 die zweite Phase begonnen, das von Ex Libris so bezeichnete „Enriched Trial“ in einer nach BVB-spezifischen Anforderungen konfigurierten Alma Sandbox.

Parallel wurden als Alternativen im Bereich Open Source „Koha“ und „Folio“ einer näheren Betrachtung unterzogen.

Bibliotheksmanagementsysteme NRW (BMS NRW)

Das Land Nordrhein-Westfalen plant bis 2021 eine neue bibliothekarische Infrastruktur einzuführen. Dazu stellt das Ministerium für Kultur und Wissenschaft (MKW) eine fünfjährige finanzielle Förderung zur Verfügung, sodass 43 Einrichtungen (Bibliotheken und hzb) auf eine einheitliche Software umsteigen können. Die in diesem Zusammenhang stehenden Arbeiten zum Aufbau eines neuen BMS NRW verlaufen fristgerecht.

Parallel sichtet das Konsortium derzeit die für die Migration noch notwendigen Arbeiten und wird dazu eigene Arbeitsgruppen gründen, die die entsprechenden Fragestellungen im Rahmen der Migration bearbeiten sollen.

BSZ-GBV-Kooperation: K10plus

Ende Juli 2017 wurde die Definition des gemeinsamen Katalogisierungsformates für K10plus, der gemeinsamen Datenbank von BSZ und GBV, abgeschlossen und parallel dazu erfolgte die Konfiguration des Titeldatenformates in der Testdatenbank, die gemeinsam von VZG und BSZ durchgeführt wurde.

An der Konfiguration der Normdaten, die nicht im GND-Format enthalten sind, und der Konfiguration bei Exemplardaten wird derzeit noch gearbeitet. Die Definition des gemeinsamen K10plus-Formates hatte zahlreiche Änderungen gegenüber den bestehenden GBV- und SWB-Formaten zur Folge, sodass für das Laden der Daten in die gemeinsame Anwendung eine Konvertierung beider Formate erforderlich ist. Diese Arbeiten sind ebenfalls weitgehend abgeschlossen.

Nach den Formatabsprachen wurde an den gemeinsamen Vorgaben für die Indexierung gearbeitet und parallel dazu auch die Konfiguration für bestimmte Felder in der Testdatenbank vorgenommen. Die administrativen Daten, also die Daten der Teilnehmerbibliotheken aus beiden Verbundsystemen, wurden in den letzten Monaten einer intensiven Analyse und Bearbeitung unterzogen. Das Ergebnis wurde Anfang Oktober 2017 in einer Testdatenbank implementiert. Die Testdatenbank ist als synchronisiertes System mit einem Primärsystem in der VZG Göttingen und einem Spiegel im BSZ Konstanz aufgesetzt. Diese Struktur dient nicht nur der Ausfallsicherheit, sondern auch dem Andocken der verbundspezifischen Anwendungen.

Mitte Oktober 2017 tagte die gemeinsame AG K10plus Katalogisierung, die sich aus den bestehenden Katalogisierungsarbeitsgruppen zusammensetzt. Themen wie Formatänderungen, -dokumentation und die kooperativ zu erstellenden Handbücher für K10plus standen dabei im Mittelpunkt. Eine Umsetzung im Produktionssystem ist für Dezember 2017 geplant. Weitere Schritte folgen im Frühjahr 2018.

Culturegraph

Das neue Team Datenmanagement in der DNB, das die Culturegraph-Prozesse zukünftig betreuen wird, hat mittlerweile seine Arbeit aufgenommen. Derzeitige Hauptaktivitäten sind die Verbesserung und Tests der bisherigen Vergleichsalgorithmen zum Zusammenfassen einzelner Publikationen zu Werkclustern sowie ein Neuaufsetzen der technischen Infrastruktur für den Import der gelieferten Verbunddatensätze, den Datenhub und die Weboberfläche.

Im Juli konnte den Verbänden zudem eine XML-Datei aller Werkcluster zur Verfügung gestellt werden, die RVK-Notationen enthalten. Diese Ausgabe kann zukünftig, nach Abnahme der Vergleichsalgorithmen, dazu dienen, eigene Titel um Elemente der Inhaltsschließung anzureichern. Insgesamt handelt es sich um 34.201.825 Titel in 10.010.580 Clustern.

Deutsche Digitale Bibliothek (DDB)

Derzeit befinden sich alle Beteiligten in einem engen Austausch, um eine Beschlussfassung zur Verfestigung der DDB und ihrer organisatorischen Verfassung herbeizuführen.

Die Anzahl nachgewiesener Objekte in der DDB liegt aktuell bei 22,75 Mio. und hat sich somit seit April 2017 um ca. 1,75 Mio. erhöht. Ebenfalls erhöht hat sich die Zahl der Datenpartner um 23 Einrichtungen, diese liegt nun bei 358. Neu hinzugekommen sind vor allem Partner aus Archiven und Museen.

DFG-Projekt LAS:eR

Das DFG-Projekt „Electronic Resource Management System“ ist auf die Entwicklung eines bundesweit einsetzbaren, mandantenfähigen Systems unter dem Namen LAS:eR (Lizenz-Administrationssystem für eRessourcen) ausgerichtet, das das Management elektronischer Ressourcen auf lokaler, regionaler und nationaler Ebene unterstützt und die Bibliotheken von den entsprechenden Verwaltungsaufgaben entlastet. Das Projektteam, bestehend aus hbz, UB Frankfurt, UB Freiburg und VZG, setzt die Entwicklungsaufgaben für den bewilligten Zeitraum von drei Jahren mit den Mitteln der DFG um.

Der Projektplan sieht vor, dass bereits nach 18 Monaten ein Testbetrieb mit ausgewählten Pilotbibliotheken etabliert ist und LAS:eR nach Fertigstellung 2019 deutschlandweit Bibliotheken und Informationseinrichtungen zur Verfügung steht. Bis Oktober 2017 wurden innerhalb der technischen Arbeitspakete definierte Aufgaben für die Pilotversion umgesetzt, wie z.B. die Anbindung von LAS:eR an die GOKb (Global Open Knowledgebase). Daneben wird die Einspielung konsortialer Lizenzen in den Datengrundbestand und ihre Pflege im Rahmen der Eigenleistung der projektbeteiligten Konsortialführer hbz, HeBIS, UB Freiburg und VZG als laufende Aufgabe bearbeitet. Auf der Basis der zentralen Knowledgebase GOKb soll eine einheitliche Nutzung von Daten zur Lizenzverwaltung elektronischer Ressourcen auf lokaler, regionaler und nationaler Ebene ermöglicht werden.

Mit dem Ziel der Internationalisierung des Systems sollen entsprechend hinterlegte Vokabulare die Mehrsprachigkeit der Oberfläche ermöglichen. Konzeptuell wurden in Abstimmung mit projektexternen Konsortialstellen eine Liste der benötigten Lizenzmerkmale erarbeitet sowie Merkmale für die Integration von E-Books und Open-Access-Publikationen definiert, damit diese technisch in Form von Properties im Datenmodell umgesetzt und komfortabel über die Oberfläche eingetragen werden können.

Auf der kommenden GASCO-Sitzung¹ im Januar 2018 werden den dort vertretenen Konsortialstellen die Funktionalitäten der für April 2018 erwarteten Pilotversion vorgestellt.

1 Die GASCO (German, Austrian and Swiss Consortia Organisation) ist die Arbeitsgemeinschaft der Bibliothekskonsortien in der Bundesrepublik Deutschland, in Österreich und in der Schweiz. Durch Absprachen innerhalb der GASCO konnten seit 2002 eine Reihe von überregionalen bzw. bundesweiten Lizenzverträgen abgeschlossen werden.

OLE/Folio

Seit 2015 wird die unter dem Dach der Quali Foundation erstellte Open-Source-Software OLE (Open Library Environment) in drei Bibliotheken produktiv eingesetzt. In den letzten Monaten wurde sowohl die Entwicklung der Systemplattform OKAPI als auch die der funktionalen Module weiter vorangetrieben.

Auf einer Webseite² informieren die deutschen Partner hbz und VZG laufend über das FOLIO-Projekt. „FOLIO“ (The Future of Libraries is Open) ist ein Next Generation Bibliothekssystem, das als Open Source entwickelt wird. Die Vorabversion des gemeinsam erarbeiteten Anforderungskatalogs für Bibliotheksmanagementsysteme der neuen Generation findet man ebenfalls auf der Webseite.

Swiss Library Service Platform (SLSP)

Das Projekt „Swiss Library Service Platform (SLSP)“ befindet sich seit März 2017 in der Aufbauphase. Im Mai 2017 wurde dazu die Aktiengesellschaft „SLSP Swiss Library Service Platform AG“ gegründet.

Auf der ersten Aktionärsversammlung im September 2017 wurde der Verwaltungsrat gewählt und die Geschäftsleitung berufen. Sieben Mitglieder des Verwaltungsrates stammen aus dem Kreis der Aktionäre. Da das achte Mitglied nicht gleichzeitig Aktionärsversammlungsmitglied und Verwaltungsrat sein soll (es soll eine bibliothekarische Fachperson sein, die aktiv in der Bibliothekswelt verankert ist), wurde im Juni 2017 Patrick Furrer zum neuen Präsidenten des Steuerungsgremiums SLSP gewählt, das Vizepräsidium behält weiterhin Susanna Bliggenstorfer. Die Ausschreibung für ein neues Bibliothekssystem ist erfolgt.

Aktuelle Information zu SLSP finden sich auf dem Projektblog unter <https://blogs.ethz.ch/slsp/>.

Strategische Überlegungen zur Zusammenarbeit der Verbundsysteme

Auf der Herbstsitzung im November 2017 der Sektion 4 (Wissenschaftliche Universalbibliotheken) des Deutschen Bibliotheksverbands e.V. (dbv) in Regensburg wurde der Vortrag von Frau Schomburg „Neue Angebote für die Wissenschaft“ zu strategischen Überlegungen zur Zusammenarbeit der Verbundsysteme mit großem Interesse aufgenommen. Dieser soll auch veröffentlicht werden.³

GVI

Der Gemeinsame Verbände-Index (GVI) ist ein Projekt der AG Verbundsysteme und wird von allen deutschen Verbänden gemeinsam weiterentwickelt. Er umfasst die Bestände aller sechs Bibliotheksverbände in Deutschland. Im Juli 2016 haben das BSZ und der KOBV zeitgleich den Beta-Betrieb des Gemeinsamen Verbände-Index aufgenommen.

² <https://www.folio-bib.org/>

³ Der Text ist im aktuellen Heft von o-bib veröffentlicht, S. 229–233. <https://doi.org/10.5282/o-bib/2018H2S229-233>.

Zwischenzeitlich sind auch die Daten der DNB und der ZDB integriert. Die Daten werden regelmäßig aktualisiert. Der technische Betrieb der unterliegenden Solr-Cloud-Anwendung wird derzeit vom BSZ übernommen. Gegenüber den bisher eingesetzten Verfahren (Metasuche über Z39.50) ist die Suche signifikant schneller, außerdem können Benutzerinnen und Benutzer Ergebnisse nun wie gewohnt über Facetten filtern. Mit einer in der HeBIS-Verbundzentrale entwickelten Lösung werden jetzt auch Normdaten ausgewertet.

Standardisierungsarbeit im deutschsprachigen Raum

Committee on Standards

Das Committee on Standards unterstützt das IFLA Governing Board und die hauptamtlich arbeitenden Organe der IFLA in allen Fragen zu den IFLA-Standards.

Durch die neue Organisationsstruktur der IFLA und die veränderte Zuordnung von Arbeitsgruppen z.B. zum Committee on Standards müssen Abläufe neu definiert werden. Hierzu soll in Kürze ein Implementation Plan for Standards in Absprache mit den Chairs der Cataloguing Section und der Review Groups ausgearbeitet werden. Im Rahmen der Global Vision der IFLA wird das Committee on Standards eine Stellungnahme abgeben. Die Gesamtergebnisse sollen im Frühjahr 2018 veröffentlicht werden.

RDA

Das vom RDA Steering Committee (RSC) im November 2016 begonnene Restrukturierungsprojekt des Standards RDA und des RDA Toolkit wurde im Rahmen der Arbeitsplanung fortgeführt. Im Juni 2018 wird das RDA Toolkit ein Relaunch erfahren und mit einer neuen Oberfläche erscheinen.

Parallel zu den Vorbereitungen für die technische Umsetzung wurden erste Entwürfe für die Restrukturierung des Inhalts erstellt, die während des RSC-Meetings im Mai in Chicago diskutiert wurden. Im Laufe der folgenden Monate wurden die Entwürfe neu ausgearbeitet und von den Mitgliedern des RSC kommentiert. Für Europa ist die EURIG und hier speziell das Editorial Committee für die Kommentierung zuständig, aus dem deutschsprachigen Raum erfolgt die Arbeit in der Fachgruppe Erschließung.

Die Anwender in Europa gehen davon aus, dass alle Arbeitsunterlagen und Schulungsmaterialien überprüft und ggf. angepasst werden müssen.

Formal wird sich das RDA Toolkit entscheidend verändern. Anstatt der bislang eher statischen Präsentation, ähnlich einer großen PDF-Datei, wird die neue Anwendung nun eine Web-Anwendung sein. Entscheidendes Instrument für die Arbeit wird die Suchfunktion sein. Folglich wird es auch keine Nummerierung der Abschnitte und kein Inhaltsverzeichnis mehr geben. Ebenso wird auf Anhänge verzichtet; die entsprechenden Inhalte werden beim zutreffenden Sachverhalt eingebracht.

GND

Der GND-Ausschuss hat sich Anfang Januar 2017 konstituiert. Eine Kooperationsvereinbarung für die GND inklusive der zur Vereinbarung gehörenden Leitlinien wurde erarbeitet und im September 2017 von allen Partnern unterzeichnet.

Als planerischer Rahmen für die Modernisierung und Öffnung der GND wurde das GND-Entwicklungsprogramm 2017–2021 veröffentlicht, das turnusgemäß aktualisiert und fortgeschrieben wird.

Parallel dazu begannen konzeptionelle Vorarbeiten zum Ausbau der Infrastruktur (Projekte „Arachne“). Weiterhin wurden zahlreiche Aktivitäten gestartet, um tragfähige Kooperationen mit den unterschiedlichen interessierten Sparten zu initiieren – u.a. im Bereich Forschung und Lehre (Projekt ORCID-DE) sowie mit Museen, Archiven, der Verlagsbranche (MVB Marketing und Verlagsservice des Buchhandels GmbH) und der Wikimedia Foundation.

BIBFRAME

Im September 2017 trafen sich in der Deutschen Nationalbibliothek in Frankfurt 40 Kolleginnen und Kollegen aus 16 europäischen Ländern und den USA zum „European BIBFRAME Workshop 2017“.

Es wurde aus sehr unterschiedlichen Perspektiven über die Ansätze und Aktivitäten berichtet. Häufig genannt wurde das Spannungsverhältnis zwischen RDA einerseits, mit dem dahinterstehenden FRBR- bzw. jetzt IFLA-Library-Reference-Modell und seiner vierstufigen Aufgliederung von Ressourcen in Werke, Expressionen, Manifestationen und Exemplare, und dem einfacheren Modell von BIBFRAME andererseits, das in der Version 2.0 für bibliografische Entitäten die dreistufige Differenzierung in Work, Instance und Item vorsieht. Hier müssen alle Beteiligten zu Klärungen kommen.

Als Fortsetzung wird es den „European BIBFRAME Workshop 2018“ geben, der voraussichtlich Mitte September 2018 auf Einladung von Casalini Libri in Florenz, Italien, stattfinden wird. Nähere Informationen (inkl. aller Vortrags-Folien, Bilder, Papiere der Hersteller und weiterführender Links) sind im Wiki-Bereich der Deutschen Nationalbibliothek unter <https://wiki.dnb.de/display/EBW> zu finden.

AG Leihverkehr

Ein Thema der Dezember-Sitzung 2017 der AG Leihverkehr war u.a. das vom Deutschen Bundestag verabschiedete Gesetz zum Urheberrecht (Urheberrechts-Wissenschaftsgesellschafts-Gesetz, UrhWissG) vom 30. Juni 2017, das am 1. März 2018 in Kraft treten wird. Es reformiert die Regelungen zur Nutzung urheberrechtlich geschützter Werke für Bildung und Forschung (sogenannte Schrankenregelungen). Welche Konsequenzen sich daraus für die Bibliotheken ergeben werden, war u. a. ein Gegenstand der Sitzung.

AG Kooperative Verbundanwendung

Die Arbeitsschwerpunkte im Berichtszeitraum waren insbesondere Fragen zur maschinellen Anreicherung von Altdaten mit RDA-Elementen, Culturegraph, Datentausch und Alma.

Bei der maschinellen Anreicherung von Altdaten mit RDA-Elementen sind der Stand der Vorbereitungsarbeiten und die Zeitplanungen in den Verbundsystemen und bei der DNB unterschiedlich, da es derzeit zwei verschiedene Herangehensweisen gibt: Zum einen, die Anreicherungselemente so genau wie möglich aus den Ursprungsdaten abzuleiten und im Zweifelsfall auf die Anreicherung zu verzichten, wenn eine genaue Ableitung nicht möglich ist; zum anderen, die Anreicherungselemente so genau wie möglich aus den Ursprungsdaten abzuleiten, zugleich aber eine flächendeckende Belegung der wichtigsten RDA-Elemente (IMD-Felder, Beziehungskennzeichen) anzustreben und dabei auch Unschärfen in Kauf zu nehmen.

Regelmäßig halbjährlich werden Datenlieferungen der Verbünde für Culturegraph bei der DNB zur Verfügung gestellt. Als mögliches weiteres Projekt neben der Anreicherung von Katalogisaten mit Sacherschließungselementen soll geprüft werden, ob eine Ergänzung von Normdatenverknüpfungen über Culturegraph erzielt werden kann.

Zeitschriftenkatalogisierung

Nach weiteren Arbeiten am neuen ZDB-Katalog wurde dieser im September 2017 in den Echtbetrieb übernommen. Er ist unter <https://zdb-katalog.de/index.xhtml> erreichbar.

Die zuletzt durchgeführten Arbeiten betrafen vor allem die auf der Syntax der Contextual Query Language basierende Expertensuche, Such- und Präsentationsmöglichkeiten für Titel in nicht-lateinischen Schriften sowie weitere neue und verbesserte Visualisierungen und Recherchemöglichkeiten. Sämtliche Suchergebnisse einschließlich spezifischer Visualisierungen sind im Produktivsystem durch stabile URLs zur weiteren Verlinkung geeignet. Details und Erläuterungen zu den verfügbaren Funktionen sind unter <https://zdb-katalog.de/help.xhtml> abrufbar. Der ZDB-Katalog wurde für mobile Endgeräte optimiert und kann damit auch „unterwegs“ bequem genutzt werden.

Die nächste Sitzung der Arbeitsgemeinschaft der Verbundsysteme findet Mitte April 2018 auf Einladung des BSZ in Konstanz statt.

Edith Röschlau, Deutsche Nationalbibliothek

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S222-228>

Neue Verbundangebote für die Wissenschaft

Bericht der Arbeitsgemeinschaft der Verbundsysteme über ihre Planungen

1. Die Arbeitsgemeinschaft der Verbundsysteme

Seit ihrer Gründung im Jahr 1983 kooperiert die Arbeitsgemeinschaft der Verbundsysteme (AGV) bei der Einführung und dem gemeinsamen Betrieb innovativer und etablierter Dienstleistungen. Die Mitglieder der AGV treffen sich in der Regel zweimal jährlich zu einem umfassenden Erfahrungsaustausch. Hierbei vereinbaren sie auch Standards für die Datenkommunikation und koordinieren die einheitliche Anwendung bibliothekarischer Regeln für die kooperative Katalogisierung, den Leihverkehr und den Datenaustausch. Ebenso werden zukünftige Ausrichtungen der Verbundangebote in diesem Forum entwickelt und besprochen.

Mitglieder der AGV sind die Verbundzentralen der Bibliotheksverbände in Deutschland, Österreich und der deutschsprachigen Schweiz sowie die Deutsche Nationalbibliothek (DNB) und die Zeitschriftendatenbank (ZDB). Gaststatus haben darüber hinaus je eine Vertretung der Deutschen Forschungsgemeinschaft (DFG) und der Kultusministerkonferenz (KMK).¹

2. Strategien und Entwicklungsziele

In einem Strategieprozess identifizierte die AGV wichtige Themen der künftigen Kooperation. Dabei orientierte sie sich an den Empfehlungen von Wissenschaftsrat und DFG hinsichtlich des Auf- und Ausbaus kooperativ gestalteter Kompetenzzentren.² Im Dialog mit Bibliotheken und Unterhaltsträgern müssen die sich daraus ergebenden Arbeitsfelder weiter ausgearbeitet und priorisiert werden. Als Informationsdienstleister sind Verbundzentralen auf eine verlässliche und kooperative Bedarfsermittlung angewiesen, in deren Kontext die organisatorischen und finanziellen Rahmenbedingungen kooperativ getragener Innovationen bestimmt werden müssen.

Zu den nach wie vor relevanten Aufgaben der Verbundzentralen, der DNB und der ZDB (im Bereich Metadatenmanagement, Standardisierung, Hosting von bibliotheksspezifischen Applikationen, Informationsversorgung wie bspw. Fernleihe) werden neue Dienste hinzukommen (z.B. publikationsnahe Dienste, Services und Infrastrukturangebote für Open Science und die Digitalisierung von Kulturbeständen).

- 1 Über die Sitzungen und Ergebnisse der Arbeitsgemeinschaft der Verbundsysteme wird regelmäßig auf der Webseite der AG der Verbundsysteme (http://www.dnb.de/DE/Wir/Kooperation/AGVerbundsysteme/agverbund_node.html) sowie bisher in der Zeitschrift „Bibliotheksdienst“ und ab 2018 in „o-bib“ berichtet.
- 2 Die Empfehlungen vom Rat für Informationsinfrastrukturen, des Wissenschaftsrats und der Deutschen Forschungsgemeinschaft zielen darauf, die Modernisierung und die Digitalisierung der Informationsinfrastrukturen für Wissenschaft und Forschung voranzutreiben und zu unterstützen. Mit den Kompetenzzentren verbindet sich die Idee der funktionalen Aufgabenteilung von Services und Dienstleistungen sowie der Nutzung von Synergien für weitere Innovationsprozesse im Bereich der wissenschaftlichen Informationsversorgung. Die Sektion 4 des dbv hat diese Empfehlungen begrüßt und in eigenen Stellungnahmen bekräftigt.

Notwendig sind zudem neue Strukturen für das Management von E-Medien inklusive Lizenzmanagement. Mit dem begonnenen Aufbau einer Infrastruktur für das Nationale Hosting elektronischer Ressourcen, dem zentralen Metadaten- und Lizenzmanagement für die Nationallizenzen und dem nationalen Statistikserver sind hier bereits wichtige Meilensteine erreicht. Derzeit erfolgt der Aufbau von LAS:eR, einem bundesweit einsetzbaren mandantenfähigen System zur Unterstützung von Bibliotheken und Konsortialstellen beim Management elektronischer Ressourcen. Auch die Unterstützung im Umgang mit Forschungsdaten und die Langzeitarchivierung von digitalen Daten mit dem Ziel der langfristigen Verfügbarkeit sind Aufgaben, denen sich die Mitglieder der AGV stellen.

Nach den durchaus kritischen Stellungnahmen der Wissenschaftsorganisationen hatten alle Verbundzentralen Anstrengungen zur Effizienzsteigerung in der Versorgung mit Katalogisierungsdiensten unternommen.

Die deutschen Verbünde möchten gemeinsam und in Kooperation mit der DNB und der ZDB zwei strategisch wichtige Themen aufgreifen:

- Die Verbesserung von Discovery- und Recherche-Diensten
- Die Aufbereitung und Bereitstellung von Inhalten für Verfahren des Text- und Data-Mining

Dadurch sollen Basisangebote entstehen, die für jede wissenschaftliche Einrichtung in Deutschland abrufbar sind und nachgenutzt werden können.

Weitere Themenfelder sind in der Diskussion und werden sich im Kontext der aktuell geführten Diskussionen um die künftige Ausrichtung des Bibliothekswesens ergeben.

3. Datendienste für Recherche-Systeme für die Wissenschaft

3.1. Ausgangslage

In den Verbunddatenbanken ist neben dem gedruckten monografischen Bibliotheksbestand bereits eine große Zahl von Nachweisen für andere Materialien wie E-Books, Mikroformen, Zeitschriften, E-Journals, Karten, Musikalien usw. und für unselbstständige Werke enthalten. Diese Nachweise bieten ein gutes Fundament zum Aufbau von bedarfsgerechten Recherchediensten für alle Medienarten.

Allerdings sind beim Nachweis von unselbstständig erschienener Literatur und Open-Access-Publikationen sowie von E-Book-Paketen kommerzieller Anbieter noch Lücken vorhanden, die mit Blick auf eine umfassende Informationsversorgung, wie sie von Wissenschaft und Forschung benötigt und eingefordert wird, geschlossen werden sollten. Hier machen Discovery-Dienste wie Summon und Primo Central von ProQuest/ExLibris, EDS von EBSCO mit riesigen globalen und aggregierten Datenbanken (Megaindices) kommerzielle Angebote. Diese können ganz oder in Ausschnitten genutzt werden und müssen, um die tatsächliche Verfügbarkeit vor Ort zu überprüfen, in einem weiteren Schritt noch mit Lizenzinformationen verknüpft werden. Trotz ihrer Größe sind jedoch auch diese Megaindices nicht vollständig und darüber hinaus in der Regel auf den angloamerikanischen Raum hin ausgerichtet.

Ergänzend und gegebenenfalls auch alternativ zu den kommerziellen Angeboten sind in den Verbänden Discovery-Indices entstanden, die neben Verbunddaten auch eine große Zahl relevanter Metadaten wie z.B. Nationallizenzen, CrossRef, Swets Online Contents, DOAJ, Medline, Springer E-Journals etc. enthalten und Bestands- bzw. Lizenzinformationen bereits auf Datenebene verbinden. Für Lizenzinformationen zu E-Journals kann dafür auf die Basisdienste der EZB und ZDB zurückgegriffen werden. Die deutschen Verbände, die ZDB und die DNB verfügen also im Bereich Discovery und Metadatenmanagement über sehr viel Erfahrung, Expertise und bereits gut genutzte Services.

3.2. Arbeitsfelder und Aufgaben

Das Ziel der AGV auf diesem Gebiet ist der Aufbau einer gemeinsamen Infrastruktur, in der Daten für Recherchesysteme in einem kooperativen, verteilten Ansatz nach einheitlichen Kriterien erschlossen und angereichert werden.

Die so aufbereiteten und verfügbaren Daten können dann wahlweise

- in kommerzielle Discovery-Plattformen,
- in verbundübergreifende Indices
- oder in individuelle Recherche-Plattformen eingespielt werden.

Damit werden die Verbände mit der DNB und der ZDB einen qualitativ hochwertigen Hintergrunddienst anbieten, der Discovery- und Recherche-Plattformen unterstützt, ohne die eine oder andere Lösung zu favorisieren.

Ergänzend zu diesem Dienst unterstützen die Verbände die von ihnen betreuten wissenschaftlichen Bibliotheken mit eigenen Recherche-Angeboten. Mit diesen zusätzlichen Dienstleistungs- und Hosting-Angeboten leisten Verbundzentralen einen wichtigen Beitrag zur allgemeinen und fachspezifischen Verfügbarkeit moderner Discovery- und Recherche-Dienste.

4. Textkorpora als Grundlage für digitale Wissenschaften

4.1. Ausgangslage

In einem interdisziplinären Forschungsumfeld nutzten die Digital Humanities ebenso wie die Sozialwissenschaften digitale Technologien zur Erforschung genuin geistes- und kulturwissenschaftlicher Sachverhalte. Dabei sind die Wissenschaftlerinnen und Wissenschaftler auf die (rechtlich) freie Verfügbarkeit von Textkorpora angewiesen, die sie mit Algorithmen des Text Data Mining analysieren können.

Die wissenschaftlichen Bibliotheken haben auf diese Anforderung reagiert und – angelehnt an die DFG-Richtlinien zur Vertragsgestaltung für Allianz-Lizenzen – damit begonnen, für Volltextangebote auch TDM-Rechte (Text-Data-Mining-Rechte) zu verhandeln. Allerdings werden diese Rechte in der Praxis oftmals nicht wahrgenommen und die Inhalte nicht für Mehrwertdienste aufbereitet. Die Gründe dafür liegen in aller Regel sowohl in mangelnden personellen als auch in mangelnden technischen Ressourcen. Die deutschen Verbände, die DNB und die ZDB können hier mit ihrer Kompetenz im Bereich Datenmanagement und ihrer langjährigen technischen Expertise attraktive

Angebote machen. Die mit dem neuen Urheberrechts-Wissensgesellschaftsgesetz 2018 in Kraft getretenen Regelungen für das Text- und Data-Mining erleichtern diese Aufgabe und machen sie zugleich dringlicher.

4.2. Arbeitsfelder und Aufgaben

Attraktive Mehrwertdienste für die Wissenschaft setzen Dienste für Text- und Data-Mining und die Bereitstellung von Korpora über standardisierte Services und Schnittstellen voraus. Diese können durch die Verbundzentralen in Kooperation mit der DNB bereitgestellt werden und damit die Angebote weiterer wissenschaftlicher Einrichtungen wie z.B. der zentralen Fachbibliotheken ergänzen. Auf Grund der sehr heterogenen Struktur der Inhalte können Datenhaltung und -aufbereitung ohnehin nur arbeitsteilig in einem Netzwerk von Partnern erfolgen.

Die Partner (Verbünde, DNB, ZDB) verfolgen das gemeinsame Ziel, der Forschung über standardisierte Schnittstellen einen rechtssicheren Zugriff gleichermaßen auf einzelne wie auch auf mehrere verteilte Korpora zu ermöglichen. Hierzu werden in enger Abstimmung TDM-Dienste mit einheitlichen Schnittstellen aufgebaut. Als Grundlage für eine offene, nachvollziehbare Wissenschaft werden diese Textkorpora durch die Vergabe von persistenten Identifikatoren zitierfähig gemacht.

In einem ersten Schritt soll mit einer Bedarfserhebung festgestellt werden, welche Forschergruppen an welchen Textkorpora interessiert sind. Hierbei werden die Universitäten, Hochschulen und Forschungseinrichtungen insbesondere in den jeweiligen (regionalen) Zuständigkeitsbereichen der Verbundzentralen einbezogen werden. Gemeinsam mit der Wissenschaft sollen notwendige Funktionalitäten aufgebaut werden – zunächst Basisservices, die disziplinübergreifend nutzbar sind. In diesem ersten Schritt geht es um die Identifikation von Quellen und den Zugriff auf die Korpora.

Der Aufbau derartiger TDM-Dienste kann ggf. in Initialprojekten durch die DFG gefördert werden. Perspektivisch wollen die Verbundzentralen mit der DNB und der ZDB in verteilter Verantwortung gemeinschaftlich TDM-Dienste mit einheitlichen Schnittstellen anbieten.

5. Kooperationen

Sowohl im Hinblick auf die Datendienste für Recherche-Systeme als auch hinsichtlich der Zugänglichkeit von Textkorpora sind die jeweiligen Kooperationspartner offen für eine Zusammenarbeit mit weiteren Akteuren aus dem deutschsprachigen Raum.

Grundsätzlich stellt die Digitalisierung der Forschung und der wissenschaftlichen Informationsinfrastrukturen tradierte Arbeitsaufteilungen und Rollen bei der Unterstützung von Wissenschaft und Forschung in Frage. Es genügt nicht, die für die analoge Welt entwickelten Modelle bezüglich Technik und Organisation in die digitale Welt zu übertragen. Digitales Forschungsdatenmanagement erstreckt sich beispielsweise über den gesamten Life-Cycle der Datengewinnung, Analyse, Publikation, Zugänglichkeit und Archivierung. In diesem Zusammenhang müssen deshalb alle Beteiligten – Forscherinnen und Forscher, Bibliothekarinnen und Bibliothekare oder IT-Expertinnen und -Experten – gemeinsam ein neues Rollenverständnis ihrer Aufgaben im digitalen Zeitalter finden.

Die Verbundzentralen zusammen mit der DNB und der ZDB verfügen über vier Jahrzehnte Organisationserfahrung, erprobte Expertise im Metadatenmanagement und über den erklärten Willen, sich aktiv in die Neugestaltung der Informationsinfrastrukturen in Deutschland einzubringen.

In den nächsten drei bis fünf Jahren sollen mit dem Aufbau von kooperativen Infrastrukturlösungen zu den beiden skizzierten Aufgabenstellungen „Datendienste“ und „Textkorpora“ durch die Mitglieder der Arbeitsgemeinschaft der Verbundsysteme wichtige Bausteine zur Versorgung von Wissenschaft, Lehre und Forschung erstellt werden. Damit nehmen die Verbünde auch die Anregungen von DFG und Wissenschaftsrat auf, sich mit neuen und kooperativen Dienstleistungen nicht nur in den Dienst von Bibliotheken, sondern auch von Wissenschaft und Forschung zu stellen.

Arbeitsgemeinschaft der Verbundsysteme

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S229-233>

Vorstand und Vereinsausschuss

Kathrin Drechsel als Vorsitzende des Regionalverbands Sachsen – Sachsen-Anhalt – Thüringen verabschiedet

Kathrin Drechsel ist aus dem Vorstand des Landesverbands Sachsen – Sachsen-Anhalt – Thüringen ausgeschieden, dem sie 12 Jahre lang in verschiedenen Funktionen angehörte, seit 2012 als Vorsitzende. In ihrer Zeit als Vorsitzende konnte sie im Jahr 2014 das 20-jährige Bestehen des VDB-Regionalverbands Sachsen – Sachsen-Anhalt – Thüringen mit einem Festakt und Kolloquium begehen. Weitere Fortbildungen zum Thema „Linked Open Data“, „Umstieg auf RVK“, „Umsetzung der EU-Datenschutzgrundverordnung in Deutschland“ sowie eine Studienfahrt folgten. 2015 beteiligte Kathrin Drechsel sich aktiv an der Diskussion und am VDB-Round-Table zum Thema „Qualifikation als wissenschaftliche Bibliothekarin/wissenschaftlicher Bibliothekar“. Die Entwicklung des Berufsfelds stand immer wieder im Fokus ihrer Arbeit.

Der Vorstand des VDB dankt Kathrin Drechsel für die nachhaltige und erfolgreiche Arbeit im Regionalverband und wünscht ihr für die Zukunft beruflich wie privat alles Gute!

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S234>

Ausschreibung der Kommissionen des VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare e. V. zum 1.10.2018

Bewerben Sie sich bis 31.7.2018!

Der VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare schreibt seine Kommissionen zum 1.10.2018 neu aus. Erstmals ausgeschrieben wird eine Kommission für forschungsnahen Dienste, die sich an Kolleginnen und Kollegen richtet, die in die Erstellung von forschungsnahen Diensten eingebunden sind. Für alle Kommissionen sind Bewerbungen von in Ausbildung befindlichen Personen besonders willkommen.

Der Vereinsausschuss beruft aus den eingegangenen Bewerbungen geeignete Vereinsmitglieder. Die Amtszeit der Kommissionen ist auf den Zeitraum Oktober 2018 bis Juni 2021 festgelegt, um zu einem Gleichklang mit den gemeinsamen Kommissionen von VDB und dbv zu kommen und diese künftig zeitgleich ausschreiben zu können.

Haben Sie Interesse an der Mitarbeit in den Kommissionen des VDB? Dann zögern Sie nicht, die Kommissionsvorsitzenden und Mitglieder der Kommissionen zu kontaktieren und sich über die Kommissionsarbeit zu informieren.

Die folgenden Ausschreibungen sind im [VDBlog](#) veröffentlicht.

Ausschreibung der Kommission für berufliche Qualifikation des VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare e. V. zum 1.10.2018



Der VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare schreibt die Kommission für berufliche Qualifikation zum 01.10.2018 neu aus. Erwünscht sind Bewerbungen von Kolleginnen und Kollegen, die Interesse an den Themen Aus- und Fortbildung, gern auch praktische Erfahrungen in diesem Bereich, z.B. als Ausbildungsleitung, mitbringen. Insbesondere Bewerbungen von in Ausbildung befindlichen Personen sind willkommen. Der Vereinsausschuss beruft aus den eingegangenen Bewerbungen geeignete Vereinsmitglieder.

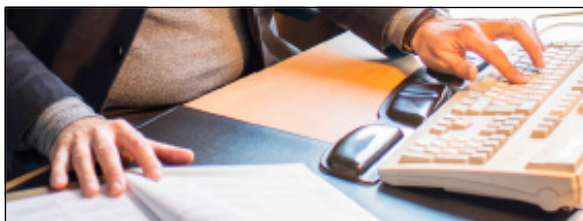
Zentrales Anliegen der Kommission für berufliche Qualifikation ist es, die Voraussetzungen und Rahmenbedingungen der bibliothekarischen Aus- und Fortbildung im Hinblick auf die Befähigung

zu lebenslangem Lernen zu thematisieren und kontinuierlich zu evaluieren sowie Stellungnahmen, Empfehlungen und Richtlinien zu erarbeiten. Die Kommission für berufliche Qualifikation bietet als regelmäßigen Service Informationen zu Ausbildung und Berufseinstieg, u.a. durch Beantwortung von Anfragen, das Website-Angebot und das jeweils beim Bibliothekartag organisierte Treffen der wissenschaftlichen Bibliothekarinnen und Bibliothekare in der Ausbildung. Ein weiteres zentrales Angebot stellt das Mentoring-Programm des VDB dar.

Haben Sie Interesse an der Mitarbeit in der Kommission für berufliche Qualifikation? Bewerbungen für die Mitgliedschaft, Amtsperiode Oktober 2018 bis Juni 2021, werden zusammen mit einem kurzen Lebenslauf und einem Motivationsschreiben (Länge ca. 200 Wörter) bis zum 31. Juli 2018 bitte nur in elektronischer Form erbeten an die:

Vorsitzende des VDB
Konstanze Söllner
Universitätsbibliothek der FAU Erlangen – Nürnberg
vorsitzende@vdb-online.org

Ausschreibung der Kommission für Fachreferatsarbeit des VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare e. V. zum 1.10.2018



Der VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare schreibt die Kommission für Fachreferatsarbeit zum 01.10.2018 neu aus. Gesucht werden Kolleginnen und Kollegen, die engagiert im Fachreferat tätig sind und Interesse an programmatischer Arbeit haben. Die Mitglieder der Kommission sollen soweit möglich die Perspektiven verschiedener Fächer, Bibliothekstypen und Regionen repräsentieren. Insbesondere Bewerbungen von in Ausbildung befindlichen Personen sind willkommen. Der Vereinsausschuss beruft aus den eingegangenen Bewerbungen geeignete Vereinsmitglieder.

Aufgabe der Kommission ist es, an der Erarbeitung und Umsetzung eines aktuellen Berufsbildes mitzuwirken, Fortbildungsveranstaltungen für Fachreferentinnen und Fachreferenten zu organisieren und Projekte für innovative Fachreferatsarbeit zu begleiten. Die Kommission hat sich darüber hinaus auch zum Ziel gesetzt, Fortbildungen mit Schwerpunkt auf den neuen Aufgabenfeldern anzubieten.

Haben Sie Interesse an der Mitarbeit in der Kommission für Fachreferatsarbeit? Bewerbungen für die Mitgliedschaft, Amtsperiode Oktober 2018 bis Juni 2021, werden zusammen mit einem kurzen

Lebenslauf und einem Motivationsschreiben (Länge ca. 200 Wörter) bis zum 31. Juli 2018 bitte nur in elektronischer Form erbeten an die:

Vorsitzende des VDB
Konstanze Söllner
Universitätsbibliothek der FAU Erlangen – Nürnberg
vorsitzende@vdb-online.org

Ausschreibung der Kommission für Rechtsfragen des VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare e. V. zum 1.10.2018



Der VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare schreibt die Kommission für Rechtsfragen zum 01.10.2018 neu aus. Erwünscht sind Bewerbungen von Kolleginnen und Kollegen, die Rechtswissenschaft studiert haben und sich für die Mitglieder des VDB engagieren möchten. Die Kommission für Rechtsfragen berät den Vereinsvorstand juristisch und erteilt den Mitgliedern des VDB – soweit rechtlich zulässig – Auskünfte in personal- und arbeitsrechtlicher Hinsicht. Insbesondere Bewerbungen von Personen in Ausbildung befindlichen Personen sind willkommen. Der Vereinsausschuss beruft aus den eingegangenen Bewerbungen geeignete Vereinsmitglieder.

Die seit 1954 bestehende Kommission für Rechtsfragen arbeitet eng mit der Rechtskommission des DBV (zuständig für institutionelle Fragen des Bibliothekswesens) zusammen und ist mit dieser immer wieder auch teilweise in Personalunion tätig.

Haben Sie Interesse an der Mitarbeit in der Kommission für Rechtsfragen? Bewerbungen für die Mitgliedschaft, Amtsperiode Oktober 2018 bis Juni 2021, werden zusammen mit einem kurzen Lebenslauf und einem Motivationsschreiben (Länge ca. 200 Wörter) bis zum 31. Juli 2018 bitte nur in elektronischer Form erbeten an die:

Vorsitzende des VDB
Konstanze Söllner
Universitätsbibliothek der FAU Erlangen – Nürnberg
vorsitzende@vdb-online.org

Ausschreibung der neuen Kommission für forschungsnahe Dienste des VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare e. V. zum 1.10.2018



Der VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare schreibt die Kommission für forschungsnahe Dienste zum 01.10.2018 erstmals aus. Erwünscht sind Bewerbungen von Kolleginnen und Kollegen, die in die Erstellung von forschungsnahe Dienstleistungen eingebunden sind. Insbesondere Bewerbungen von in Ausbildung befindlichen Personen sind willkommen. Der Vereinsausschuss beruft aus den eingegangenen Bewerbungen geeignete Vereinsmitglieder.

Forschungsnahe Dienste werden von Bibliotheken in immer größerem Maße wahrgenommen. Weitere neue Arbeitsschwerpunkte entstehen durch die Transformation hin zu publikationsbasierten Finanzierungsstrukturen. Dies wird auch durch die Zahl der Stellenausschreibungen deutlich. Zu diesem Aufgabenspektrum gehören: Open Access und Publikationsdienstleistungen, Forschungsdatenmanagement und Forschungsinformationssysteme, Bibliometrieservices sowie weitere Aufgaben im Bereich von Wissenschaftsmanagement und Open Science, insbesondere auch für die Digital Humanities.

Diese Arbeitsbereiche wachsen mit zunehmend digitaler Literatur- und Informationsversorgung und sind häufig durch eine projektförmige Finanzierung, ein umfangreiches Antragsmanagement und eine Querschnittsstellung gegenüber dem disziplinär ausgerichteten Fachreferat charakterisiert. Ein weiterer Unterschied gegenüber dem stärker im Bereich der Vermittlung von Services und Informationskompetenz angesiedelten Fachreferat besteht im konkreten Erstellen der Dienstleistung. Häufig bestehen bei der Erstellung forschungsnahe Dienste enge Kooperationen mit den IT-Abteilungen der Bibliotheken, den Forschungsabteilungen der Hochschulen und über die Grenzen der eigenen Einrichtung hinaus.

Die Erstellung von neuen Dienstleistungen für die Wissenschaft ist in der Arbeit des VDB derzeit noch zu wenig berücksichtigt. Kolleginnen und Kollegen, die stark in die Erstellung von forschungsnahe Dienstleistungen eingebunden sind, sollen deshalb künftig durch entsprechende Fortbildungs- und Vernetzungsangebote des VDB noch besser unterstützt werden. Vorstand und Vereinsausschuss haben daher beschlossen, dieses berufs- und laufbahnübergreifende Thema mit einer neuen Kommission zu institutionalisieren. Zu diesem Zweck wird eine „Kommission für forschungsnahe Dienste“ ausgeschrieben.

Die Aufgaben der „Kommission für forschungsnahе Dienste“ sollen zunächst darin liegen, geeignete Angebote für die VDB-Mitglieder sowie die passenden Arbeits- und Fortbildungsformate zu entwickeln. Darüber hinaus soll die Kommission mit eigenen Stellungnahmen und Konzepten die Entwicklung des Berufsfelds begleiten und mitgestalten.

Haben Sie Interesse an der Mitarbeit in der Kommission für forschungsnahе Dienste? Bewerbungen für die Mitgliedschaft, Amtsperiode Oktober 2018 bis Juni 2021, werden zusammen mit einem kurzen Lebenslauf und einem Motivationsschreiben (Länge ca. 200 Wörter) bis zum 31. Juli 2018 bitte nur in elektronischer Form erbeten an die:

Vorsitzende des VDB
Konstanze Söllner
Universitätsbibliothek der FAU Erlangen – Nürnberg
vorsitzende@vdb-online.org

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S235-239>

Gemeinsame Kommissionen von VDB und dbv – Mitglieder der neuen Amtszeit

Während die Ausschreibung der VDB-Kommissionen noch läuft (Bewerbungsfrist 31. Juli 2018), wurden die Mitglieder der Gemeinsamen Kommissionen von VDB und dbv für die neue Amtszeit ab 1. Juli 2018 bereits berufen. In den Kommissionen engagieren sich diese Kolleginnen und Kollegen:

Baukommission

- Ulrike Brunenberg-Piel, ULB Düsseldorf
- Olaf Eigenbrodt, SuUB Hamburg
- Susanne Kandler, UB Freiberg
- Oliver Kohl-Frey, UB Konstanz
- Prof. Dr. Michael Mönnich, KIT Karlsruhe
- Dr. Alice Rabeler, ULB Bonn

Kommission für Informationskompetenz

- Carolin Ahnert, TUB Chemnitz
- Claudia Martin-Konle, UB Gießen
- Dr. Oliver Schoenbeck, BIS Oldenburg
- Dr. Marcus Schröter, UB Freiburg
- Erik Senst, UB Bielefeld
- Dr. Naoka Werr, HföD München

Managementkommission

- Friederike Sablowski, StB Bad Seegeberg
- Martin Lee, FUB Berlin
- Daniela Poth, SUB Göttingen
- Prof. Cornelia Vonhof, HdM Stuttgart
- Frauke Untiedt, Bücherhallen, Hamburg
- Dr. Isabelle Tannous, Stiftung Wissenschaft und Politik, Berlin

Der Vorstand des VDB dankt allen Kommissionsmitgliedern, die aus den gemeinsamen Kommissionen ausgeschieden sind, sehr für ihr Engagement und die erfolgreiche gemeinsame Arbeit der letzten Jahre. Den neuen Kolleginnen und Kollegen sowie den Mitgliedern, die in der kommenden Amtsperiode ihre ehrenamtliche Arbeit fortsetzen, dankt der Vorstand herzlich für Ihre Bereitschaft, sich für die wichtige fachliche Arbeit einzusetzen und wünscht den gemeinsamen Kommissionen weiterhin viel Erfolg!

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S240>

Bibliothek – Qualifikation – Perspektiven: D-A-CH-S-Tagung der bibliothekarischen Verbände

13. bis 14. Februar 2019, Ludwig-Maximilians-Universität München



Die gemeinsame D-A-CH-S-Tagung (Deutschland, Österreich, Schweiz, Südtirol) steht unter dem Titel „Bibliothek – Qualifikation – Perspektiven: D-A-CH-S-Tagung der bibliothekarischen Verbände“ und findet vom 13. bis 14. Februar 2019 an der Ludwig-Maximilians-Universität in München statt. Das grenzübergreifende Thema soll der Intensivierung der Zusammenarbeit der Verbände und dem beruflichen Austausch dienen.

Tagungsprogramm

Mittwoch, 13. Februar 2019

10:00–12:00 Uhr **Nachwuchsforum, separate Ausschreibung**

11:30 Uhr **Welcome – Tagungsbüro geöffnet**

13:00–13:30 Uhr **Eröffnung der Tagung & Grußworte**

13:30–14:15 Uhr **Keynote**

- **Prof. Dr. Meike Sophia Baader** (Vizepräsidentin der Stiftung Universität Hildesheim, Institut für Erziehungswissenschaften): Bibliotheken als (Bildungs)Räume für Forschung und Lehre

14:15–15:30 Uhr **Diskussion I: Ländervergleichende Bestandsaufnahme der aktuellen Bibliotheksausbildung in D-A-CH-S**

- Podiumsworkshop mit Moderation
- 5'-Statements aus den einzelnen Ländern
- anschließende Diskussion organisiert durch **Dr. Klaus Gantert**, Vorsitzender der KIBA

15:30–16:00 Uhr **Pause**

16:00–17:30 Uhr **Vortragsblock I: Anforderungen an die Berufsqualifikation**

- Aus Sicht der Fachverbände (national / international): **Konstanze Söllner** (VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare)
- Aus Sicht der Wissenschaft: **Prof. Dr. Konrad Förstner** (ZB MED/TH Köln): Literatur, Daten, Software – Neue Aufgabenbereiche der Bibliothek
- Aus Sicht des Rates für Informationsinfrastrukturen: **Prof. Dr. Michael Jäckel** (Universität Trier): Wie schlägt das Herz der Universität? Rhythmusstörungen und Schrittmacher im digitalen Zeitalter
- Aus Sicht der Bibliotheken: **Prof. Dr. Rudolf Mumenthaler** (Zentral- und Hochschulbibliothek Luzern): Anforderungen an Bibliothekar/innen aus Sicht der Zentral- und Hochschulbibliothek Luzern

19:00 Uhr **Social Event: Typisch bayrisches Spanferkelessen im Hofbräuhaus**

Donnerstag, 14. Februar 2019

8:30 Uhr **Tagungsbüro geöffnet**

9:00–10:30 Uhr **Vortragsblock II: Entwicklungen in der Berufsausbildung 1: Inhalte, Formate, Kooperationen / Schlaglichter der Ausbildung**

- [Andrea Betschart](#) (Ausbildungsdelegation I+D): Die Berufsbildung der Schweiz, wie sie am Puls der Berufswelt und alltagstauglich bleibt
- [Susanne Tretthahn](#) (Büchereiverband Österreichs): Neue Wege – Ausbildung für hauptamtliche Bibliothekarinnen und Bibliothekare in Österreich
- [Marion Gamper](#) (Amt für Bibliotheken und Lesen Bozen): Quer rein in die Bibliothek! Ausbildung von Quereinsteiger/innen und Lehrpersonen in Südtirol
- [Franziska Zenkel/Hannah Dassler](#) (Studierende der HföD München): „Where did you come from, where did you go?“ – Das erwartet Studierende des Diplom-Studiengangs Bibliothekswesen an der Hochschule für den öffentlichen Dienst in München

10:30–11:00 Uhr **Pause**

11:00–12:30 Uhr **Vortragsblock III: Entwicklungen in der Berufsausbildung 2: Nationale / Internationale Durchlässigkeit**

- [Dr. Christian Oesterheld](#) (Zentralbibliothek Zürich): „Einmal Berlin – Zürich einfach“: Ausbildung und Berufswege zwischen Deutschland und der Schweiz – nahe und doch (ein wenig) anders. Einige (persönliche) Anmerkungen.
- [Petra Imwinkelried](#) (Kantonsbibliothek Glarus): Der kleine Gegentrend – von der Schweiz nach Deutschland
- [Eva Ramminger](#) (Universitäts- und Landesbibliothek Tirol): Bibliotheksausbildung zwischen Durchlässigkeit und Exzellenz
- [Susanne Kandler](#) (Universitätsbibliothek der TU Freiberg): Ausbildung in Österreich, Berufsweg nach Deutschland – ein kurzer Erfahrungsbericht

12:30–13:30 Uhr **Mittagspause**

13:30–15:00 Uhr **Vortragsblock IV: Entwicklungen in der Berufsausbildung 3: Perspektiven / Anforderungen / Konkurrenz**

- [Dr. Ina Blümel](#) (Hochschule Hannover): Neue Berufsbilder an Bibliotheken – Auswirkungen auf die informationswissenschaftliche Ausbildung an der Hochschule Hannover und Beispiele aus der Praxis
- [Prof. Ivo Macek](#) (HTW Chur): Die Bibliothekswelt im Wandel: Curriculare Auswirkungen aus Sicht eines Studiengangleiters am Beispiel des ‘Master of Advanced Studies in Information Science’ der HTW Chur
- [Prof. Dr. Elke Greifeneder](#) (IBI der HU Berlin): Weiterbildung als Bildung und nicht nur als Abschlussgrad: Erfahrungen aus dem Weiterbildenden Fernstudiengang
- [Prof. Magnus Pfeffer](#) (HdM Stuttgart): Informatik für alle? Wie reagieren wir in der bibliothekarischen Ausbildung auf die Herausforderungen der Informationsgesellschaft und Veränderungen im Berufsbild?

15:00–15:30 Uhr Pause

15:30–16:30 Uhr Diskussion II: Anforderungen an die Bibliotheksausbildung im 21. Jahrhundert

Diskutant/innen:

- Dr. Arne Ackermann (Stadtbibliothek München),
- Marion Gamper (Amt für Bibliotheken und Lesen Bozen),
- Dr. Caroline Leiß (TUM),
- Mag. Monika Schneider-Jakob MSc (Universitätslehrgang ULB Innsbruck),
- Herbert Staub (BIS),
- Franziska Zenkel (HföD)

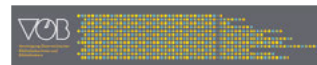
Teilnahmegebühr

- **Mitglieder der Verbände** BIB, VDB, ALBAD, ALA, BIS, BVS, BVÖ, VÖB: Frühbucher 95 EUR, ab 16.12.18: 120 EUR
- **Ermäßigte Teilnahmegebühr:** Frühbucher 45 EUR, ab 16.12.2018: 70 EUR (Gültig für Auszubildende, Studierende (auch von berufsbegleitenden Studiengängen), Teilnehmer/-innen an Fachwirtausbildungen, Referendar/-innen, Arbeitslose, Rentner/-innen)
- **Sonstige Interessierte:** Frühbucher 145 EUR, ab 16.12.18: 170 EUR

Anmeldung über die Tagungs-Website: <http://www.ub.uni-muenchen.de/DACHS-Tagung>



Berufsverband Information Bibliothek e.V.



Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S241-243>

Kommissionen

Kommission für Rechtsfragen

Tätigkeitsbericht 2017

In der Kommission für Rechtsfragen des VDB – Verein Deutscher Bibliothekarinnen und Bibliothekare arbeiten die Kollegen/inn/en Jost Lechte (Bielefeld), Markus Lohmann (Bochum), Kathrin Schwärzel (München), Verena Simon (Bonn) und Thomas Witzgall (Jena) zusammen.

Die Kommission traf sich im Berichtszeitraum zweimal: Eine Sitzung fand in Frankfurt am Main anlässlich des 106. Deutschen Bibliothekartages und eine zweite im September in Würzburg statt.

Die von den Vereinsmitgliedern gestellten Anfragen aus allen Bereichen des Arbeits- und Dienstrechts wurden wie in den vergangenen Jahren vertraulich behandelt, und daher arbeitet die Kommission auch eher im Hintergrund. Erfreulicherweise konnten alle Anfragen im Berichtszeitraum telefonisch oder per E-Mail beantwortet werden.

Einen weiteren Schwerpunkt der Kommissionsarbeit bildet die Beratung des Vorstandes in juristischen Fragen. Für 2017 ist hier insbesondere die Reform des Urheberrechts zu nennen, die der VDB im Gesetzgebungsverfahren, aus Sicht der Beschäftigten in den Bibliotheken, konstruktiv kommentierte und auch weiterhin in der tatsächlichen Anwendung kritisch begleiten wird.

Ebenfalls notwendig waren Vorarbeiten und Anpassung vor Inkrafttreten der Datenschutz-Grundverordnung.

Wie in früheren Jahren hielten die Mitglieder der Rechtskommission auch im Berichtsjahr 2017 engen Kontakt zu anderen bibliothekarischen Gremien und Einrichtungen, z.B. der ajbd, der ViFa Recht, als Mitglied der AhD sowie als Vortragende / Referentinnen und Referenten im bibliothekarischen Fortbildungsbereich. Zur Rechtskommission des dbv ergaben sich die engen Kontakte naturgemäß aus der Abgrenzung der Aufgabenfelder und der Doppelmitgliedschaft zweier Mitglieder in beiden Kommissionen.

Thomas Witzgall, Thüringer Universitäts- und Landesbibliothek Jena (Vorsitzender der Kommission)

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S244>

Gemeinsame Baukommission des Deutschen Bibliotheksverbands und des Vereins Deutscher Bibliothekarinnen und Bibliothekare

Jahresbericht 2017

1. Mitglieder

Der Kommission gehören in der Amtszeit 2017 folgende Mitglieder an:

- Susanne Kandler von der Universitätsbibliothek der TU Bergakademie Freiberg (Vorsitz)
- Ulrike Brunenberg-Piel von der ULB Düsseldorf
- Olaf Eigenbrodt von der Staats- und UB Hamburg Carl von Ossietzky
- Oliver Kohl-Frey von der Universität Konstanz / KIM
- Milena Pfafferott von der UB Ilmenau
- Herr Prof. Mönnich von der KIT-Bibliothek Karlsruhe (seit Januar 2017)

2. Arbeitssitzungen

Die Baukommission hat sich seit der letzten Beiratssitzung dreimal getroffen:

- 06.03.2017 Arbeitssitzung SUB Hamburg
- 30.05.2017 Arbeitssitzung in Frankfurt am Main (Bibliothekartag)
- 17.09.2017 Arbeitssitzung UB Ilmenau

3. Veranstaltungen

Fortbildung: „Technische Aspekte der Sanierung von Bibliotheksgebäuden“

Am 6. und 7. März 2017 fand in Hamburg eine zweitägige Weiterbildung zum Thema „Technische Aspekte der Sanierung von Bibliotheksgebäuden“ statt, zu der die Baukommission 30 Teilnehmende begrüßen konnte.

Die Fortbildung befasste sich mit der Erkennung, Beschreibung und Planung notwendiger Sanierungsmaßnahmen und ihrer Abgrenzung gegenüber anderen Modernisierungen.

Bibliothekartag 2017: „Neue Raumideen in Bibliotheken: Umnutzung, Aufwertung, Erweiterung“

Bei der Invited Session auf dem 106. Bibliothekartag am 31. Mai 2017 in Frankfurt am Main wurden innovative „Neue Raumideen“ aus der Bibliothek der ZHAW Winterthur, der Stadtbücherei Oberhausen und der UB Freiburg vorgestellt.

Die Veranstaltungswebsite mit Programm, Berichten und Vortragspräsentationen sind online unter: <https://www.bibliotheksverband.de/fachgruppen/kommissionen/baukommission/fortbildung.html>.



Robert Zepf, UB Rostock, bei der Veranstaltung der Baukommission zur Evaluierung der DIN 67700 beim Bibliothekartag in Berlin 2018. Foto: Dirk Deckbar

4. Weitere Aktivitäten

DIN 67700

Die neue DIN 67700 „Bau von Bibliotheken und Archiven“ wurde von der Baukommission intensiv geprüft und es fanden weitere Gespräche statt; die Baukommission evaluierte die DIN 67700 mit Bauexpertinnen und -experten beim Bibliothekartag 2018 in Berlin.

Beratung

Das Angebot der Beratung und Unterstützung von Kolleginnen und Kollegen bei Fragen von Bibliotheksbau und -gestaltung wurde gut angenommen, es wurden unterschiedliche Fragen an die Baukommission herangetragen.

5. Ausblick

Informationsportal

Ausbau der Website zu einem Informationsportal: Ggeplant ist eine „Neuaufgabe des Bibliotheksbauarchives“, um einen Überblick über die verschiedenen Um- und Neubauten zu bieten und sie dort zu bündeln.

Susanne Kandler, Universitätsbibliothek der TU Freiberg (Vorsitzende der Kommission)

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S245-246>

Landes- und Regionalverbände

Regionalverband Berlin – Brandenburg

Jahresbericht 2017

Aktivitäten

Seit vielen Jahren organisiert der VDB-Regionalverband Berlin – Brandenburg gemeinsam mit anderen bibliothekarischen Verbänden aus der Region den „Tag der Bibliotheken in Berlin und Brandenburg“ als Angebot zur fachlichen Weiterbildung und zum Austausch für die Kolleginnen und Kollegen. Am 11. März 2017 war es wieder einmal so weit: Der nunmehr 6. Tag der Bibliotheken fand dieses Mal in Cottbus (Brandenburg) statt. Das Motto lautete „Offenheit und Integration – Menschen und Bibliotheken in Bewegung“. Neben Grußworten aus der Politik von Frau Dr. Martina Münch, Ministerin für Wissenschaft, Forschung und Kultur des Landes Brandenburg, und Herrn Prof. Dr.-Ing. Matthias Koziol, Vizepräsident für Lehre und Studium der BTU Cottbus-Senftenberg, sowie Herrn Berndt Weiß, Leiter des Geschäftsbereichs Jugend, Kultur, Soziales der Stadt Cottbus, gab es eine Eröffnungsrede von Thomas Feibel – Autor, Journalist, Medienexperte und Preisträger der Karl-Preusker-Medaille 2014. Anschließend warben die Referentinnen und Referenten bei einer „Minute Madness“ um die Gunst des Publikums. In nur 90 Sekunden stellten sie kurz und sehr kreativ das Thema vor, das sie in den kommenden Sessions intensiver diskutieren wollten. So konnten sich alle Teilnehmerinnen und Teilnehmer einen besseren Überblick über die Themen verschaffen und sich überlegen, in welche anschließenden Vorträge und Workshops sie gehen möchten. Die Bandbreite der Themen war dem Motto entsprechend groß: von Barrierefreiheit und Inklusion über Integration von Geflüchteten bis hin zu innovativen Lern- und Arbeitsräumen sowie Digitalisierung. Auch die Formate wurden flexibilisiert: Es gab Vorträge, World-Cafés, Bibliotheksführungen, Workshops u. Ä. Einen ausführlichen empfehlenswerten Tagungsbericht hat Frau Indra Heinrich, die Schatzmeisterin des VDB-Regionalverbandes Berlin und Brandenburg, verfasst.¹

Das zweite Highlight des Jahres 2017 für den VDB-Regionalverband Berlin und Brandenburg war die Vergabe des *Innovationspreises* für Bibliotheken der Länder Berlin und Brandenburg, an dessen Konzeption sich der Regionalverband bereits seit 2016 intensiv beteiligte.

Unter der Ägide der Stiftung Bibliotheksforschung e. V. wurde der Preis erstmalig 2017 von den Ausbildungseinrichtungen sowie den Bibliotheksverbänden der Region vergeben. Offiziell ausgerufen wurde er am 11. März 2017 auf dem „6. Tag der Bibliotheken in Berlin und Brandenburg“ in Cottbus. Der Preis war mit 10.000 EUR dotiert. Das Ziel des Innovationspreises ist es, mit Hilfe der finanziellen Unterstützung innovative und herausragende Ideen zur Bibliotheksarbeit in Berlin und

1 Heinrich, Indra: Tag der Bibliotheken in Berlin & Brandenburg 2017. In: *o-bib – das offene Bibliotheksjournal*, Ausgabe 2/2017, online kostenfrei abrufbar unter: <https://www.o-bib.de/article/view/2017H2S105-110/5892>. Weitere Impressionen finden Sie im Blogbeitrag bei Lesewolke: <https://lesewolke.wordpress.com/2017/03/16/tdbbb-2017/> [zuletzt geprüft am 06.06.2018] und in einem Kurzvideo auf YouTube: https://www.youtube.com/watch?v=9QS_nQQPFSQ [zuletzt geprüft am 06.06.2018].

Brandenburg zu realisieren und so zukunftsfähige Bibliotheksarbeit in der Region zu fördern. Es ist geplant, den Preis alle zwei Jahre zu verleihen.²

Die diesjährige Preisverleihung fand am 23. Oktober 2017 in der Vertretung des Landes Brandenburg beim Bund statt. Es wurden insgesamt zehn Wettbewerbsanträge fristgerecht eingereicht und lagen der Jury zur Begutachtung vor, von denen fünf für den Preis nominiert worden sind. In der Preis-Jury wurde der VDB-Regionalverband Berlin und Brandenburg von Frau Indra Heinrich vertreten. Die Preisverleihung wurde von Staatssekretär Martin Gorholt als Bevollmächtigtem des Landes Brandenburg beim Bund und Hausherrn feierlich eröffnet. Den Festvortrag hielt Prof. Dr. Jan-Hendrik Olbertz, ehemaliger Kultusminister von Sachsen-Anhalt und Präsident der Humboldt-Universität zu Berlin. Im Anschluss stellten die fünf nominierten Preisanwärter dem Publikum ihre Ideen und Konzepte für innovative Bibliotheksarbeit vor: Die Bibliothek Wirtschaft und Management der TU Berlin möchte in Form innovativer Online-Tutorials das Erlernen der Methode des Systematic Literature Review (SLR) anbieten; die Stadt- und Landesbibliothek Potsdam verbindet mit dem **Medienmacher-Bibliotheks-Mobil** auf innovative Weise dezentrale Bibliotheksarbeit mit der Vermittlung von Medienkompetenz (Gaming/Gamedesign, Trickfilm/Film, Comic, Programmieren/Coding und Podcast) und zielt darauf ab, insbesondere Schulen in sozialen Brennpunkten zu erreichen; die Stadtbibliotheken in Pankow möchten mit der Bibliothek der Dinge den gesellschaftlichen Trend der Sharing Economy aufgreifen und partizipatorische Bibliotheksarbeit gestalten; der Verbund der Öffentlichen Bibliotheken Berlins wagt mit „Filmfreund – Das Streaming-Portal für Bibliotheken“ eine Alternative zu Netflix & Co.; die Stadtbibliothek Wildau möchte zukünftig Leseförderung mit Hilfe eines handlichen humanoiden Roboters umsetzen und hatte einen entsprechenden Prototypen gleich mitgebracht.³ Am Ende überzeugte der Roboter, und die Stadtbibliothek Wildau durfte sich über den ersten Platz freuen.⁴ Die Preisübergabe wurde u.a. von Dr. Ulrike Gutheil, der Staatssekretärin für Wissenschaft, Forschung und Kultur des Landes Brandenburg, und Dr. Torsten Wöhlert, dem Staatssekretär für Kultur bei der Senatsverwaltung für Kultur und Europa, vorgenommen. Musikalisch umrahmt wurde das Programm vom „Bibliothekarischen Berlin-Brandenburgischen Streichquartett“.

Mitgliederversammlung

Die Mitgliederversammlung des VDB-Regionalverbands Berlin und Brandenburg fand turnusgemäß und fristgerecht am 23. Oktober 2017 in der Vertretung des Landes Brandenburg beim Bund ab 16:00 Uhr im Anschluss an die Preisverleihung des Innovationspreises für Bibliotheken der Länder Berlin und Brandenburg statt. Die Vorstandsvorsitzende, Frau Janin Präßler, begrüßte alle Anwesenden. Nach dem Jahresbericht und Neuigkeiten aus dem Bundesvorstand sowie dem Bericht der Schatzmeisterin wurde der Vorstand entlastet. Danach gab es Zeit für Fragen und Wünsche der Mitglieder. Diskutiert wurde u. a. die Frage, wie neue, junge Mitglieder gewonnen werden können. Außerdem wurde der Wunsch nach mehr Fortbildungsangeboten und einer besseren Sichtbarkeit der VDB-Angebote verdeutlicht. Im Ausblick wurde auf den 107. *Bibliothekartag* hingewiesen, der

2 Weitergehende Informationen: <http://bibliotheksforschung.de/innovationspreis> [zuletzt geprüft am 06.06.2018]

3 Detaillierte Projektbeschreibungen: <http://bibliotheksforschung.de/innovationspreis/nominierungen/> [zuletzt geprüft am 06.06.2018]

4 Vgl. <http://www.maz-online.de/Brandenburg/Roboter-aus-Wildau-holen-Innovationspreis> [zuletzt geprüft am 06.06.2018]

inzwischen vom 12.-15. Juni 2018 in Berlin stattgefunden hat.⁵ Der Regionalverband engagierte sich im Ortskomitee. Zudem lud er die Mitglieder herzlich zur Podiumsdiskussion „Qualifiziert, motiviert, befristet und in Teilzeit - Wie prekär ist das Bibliothekswesen?“ ein; im Anschluss gab es beim Stand der Verbände einen Empfang des Landesverbandes, der auf große Resonanz stieß.



Die Podiumsdiskussion „Qualifiziert, motiviert, befristet und in Teilzeit - Wie prekär ist das Bibliothekswesen?“ des Regionalverbands Berlin - Brandenburg beim Berliner Bibliothekartag war ein voller Erfolg. Foto: Dirk Deckbar

Anzahl der Mitglieder

Der Regionalverband Berlin-Brandenburg hat aktuell 265 Mitglieder. 2017 sind 11 neue Mitglieder eingetreten.

Janin Präßler, Stadtbibliotheken Treptow-Köpenick, Berlin (Vorsitzende des Regionalverbandes)

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S247-249>

5 Vgl. <https://bibliothekartag2018.de/>

Landesverband Hessen

Hessischer Bibliothekstag 2018, 14. Mai 2018 in Wetzlar

Der hessische Bibliothekstag, gemeinschaftlich organisiert vom Landesverband Hessen im Deutschen Bibliotheksverband (dbv) und vom VDB Landesverband Hessen, fand in diesem Jahr in Wetzlar statt. In den Räumen des Kreistags des Lahn-Dill-Kreises sowie in der im letzten Jahr am neuen Standort wiedereröffneten Stadtbibliothek Wetzlar gab es für die knapp 150 Anwesenden spannende Beiträge rund um das diesjährige Thema „Allianzen für Bildung“.

Eine Einführung in das Thema bot der Festvortrag, der in diesem Jahr von Petra Büning, Mitglied des Bundesvorstandes des dbv gehalten wurde. Unter dem Titel „Heiße Eisen und wie man sie schmiedet. Lobbyarbeit für Bibliotheken“ präsentierte Frau Büning einen interessanten und äußerst kurzweiligen Einblick in die Lobbyarbeit des dbv. Schwerpunktmäßig berichtete Frau Büning über die gelungene Lobbyarbeit des dbv zum Thema Urheberrecht für die Wissensgesellschaft.

Zuvor war bereits zum 21. Mal der Hessische Bibliothekspreis der Sparkassen-Kulturstiftung und des Landesverbandes Hessen im Deutschen Bibliotheksverband verliehen worden. Er ging dieses Jahr an die Stadtbibliothek Dreieich.



Foto: Mohamad Osman

Im letzten Vortrag des Tages, „Copyright office‘ in der Bibliothek?“ stellte Stephanie Hasler von der Goethe-Universität Frankfurt die Frage, ob sich das bereits jetzt vielfältige Tätigkeitsfeld von Bibliothekarinnen und Bibliothekaren nicht noch durch Urheberrechtsberatung weiter auffächern wird. An Beispielen aus dem anglo-amerikanischen Raum zeigte Frau Hasler, dass dortige Bibliotheken bereits über ein „Copyright Office“ bzw. einen „Copyright Officer“ verfügen. Die Vermittlung von Urheberrechts-Kompetenz wird, so die Ausführungen von Frau Hasler, zukünftig wohl eine wichtige neue Aufgabe für juristisch vorgebildete Bibliothekarinnen und Bibliothekare werden und damit den bereits bestehenden Aufgabenbereich Informationskompetenz um eine neue Facette erweitern.

Inhaltlich konzipiert, organisiert und moderiert wurde das Forum Wissenschaftliche Bibliotheken vom Vorstand des VDB-Landesverbandes Hessen.

Julijana Nadj-Guttandin, Deutsche Nationalbibliothek (Schriftführerin des Landesverbandes Hessen)

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S250-252>

Regionalverband Sachsen – Sachsen-Anhalt – Thüringen

Mitgliederversammlung: Wahl eines neuen Vorstandes

Am 07.05.2018 fand an der UB Erfurt die Mitgliederversammlung des Landesverbandes Sachsen – Sachsen-Anhalt – Thüringen statt.

Erster Programmpunkt war ein Bericht des Vorstandes; u.a. über die Aktivitäten des letzten Jahres durch die Vorsitzende Kathrin Drechsel. Es wurden zwei Veranstaltungen durchgeführt: ein Symposium zum Thema Quelle(n) für Inspiration: *20 Jahre Sammlung Teufel an der Universität Erfurt und die Veranstaltung Datenschutz: Grundlagen. Umsetzung der EU-Datenschutzgrundverordnung in Deutschland*, von Prof. Kay Gürtzig, an der UB Erfurt, welche im Vorfeld der Mitgliederversammlung stattgefunden hat. Eine dritte geplante Veranstaltung, ein Barcamp „Werkstatt Fachreferat“ im September 2017 wurde aufgrund geringen Interesses abgesagt.

Im Anschluss wurde der Kassenbericht durch den Kassenprüfer Herrn Thomas Witzgall vorgetragen und der Vorstand entlastet.

Wichtigster Programmpunkt war die Wahl eines neuen Vorstandes für die nächsten zwei Jahre.

Frau Kathrin Drechsel gab bekannt, dass sie nicht mehr für den Vorstand des Landesverbandes Sachsen – Sachsen-Anhalt – Thüringen kandidieren wird. Wir bedanken uns sehr herzlich für das außerordentliche Engagement, welches Frau Kathrin Drechsel die letzten Jahre im Rahmen ihrer Arbeit für den Vorstand des Regionalverbandes Sachsen – Sachsen-Anhalt – Thüringen gezeigt hat.

In den neuen Vorstand wurden Silke Berndsen zur neuen Vorsitzenden, Dr. Susann Özüyaman zur Stellvertreterin und Anika Richter zur neuen Schatzmeisterin gewählt. Alle drei gehörten bereits in den letzten zwei Jahren dem Vorstand an und haben nun neue Ämter innerhalb des Vorstandes übernommen.

Susann Özüyaman, ULB Sachsen-Anhalt (Stellvertretende Vorsitzende des Landesverbandes Sachsen – Sachsen-Anhalt – Thüringen)

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S253>

Landesverband Bayern

Eine Zusammenfassung der Fortbildung „Fachinformationsdienste: Angebote und Perspektiven der überregionalen Informationsversorgung“

In den letzten Jahren entstanden eng auf die jeweilige Fachcommunity ausgerichtete, aber je nach Wissenschaftsdisziplin sehr verschiedene Fachinformationsdienste. Dieser heterogenen „FID-Landschaft“ widmete sich der Landesverband Bayern in einer Fortbildung, über die Rainer Plappert, UB Erlangen-Nürnberg, in diesem Heft berichtet: **Neue Konzepte für die Informations- und Literaturversorgung? Eine Zusammenfassung der Fortbildung „Fachinformationsdienste: Angebote und Perspektiven der überregionalen Informationsversorgung“**, S. 208–213, <https://doi.org/10.5282/o-bib/2018H2S208-218>.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S254>

Aus der Mitgliederverwaltung

Berücksichtigt sind Hinweise, die die Redaktion bis Redaktionsschluss dieses Heftes erreicht haben. Hierbei kann es im Einzelfall auch zu nachträglichen Korrekturen seitens der Mitgliederverwaltung gekommen sein.

Neue Mitglieder

Wir begrüßen sehr herzlich neue Kolleginnen und Kollegen im VDB!

Aufschnaiter, Stefanie

Bayerische Staatsbibliothek München

Begrich, Uta

Universitäts- und Stadtbibliothek Köln

Bentrup, Stefanie

Staats- und Universitätsbibliothek Hamburg Carl von Ossietzky

Berberich, Dr. Stefanie

Universitätsbibliothek Heidelberg

Beringer, Constanze

Deutsche Zentralbibliothek für Medizin (ZB MED) – Informationszentrum
Lebenswissenschaften, Köln

Brauer, Julia

Thüringer Universitäts- und Landesbibliothek Jena

Breckheimer, Matthias

Leibniz-Institut für Geschichte und Kultur des östlichen Europa (GWZO), Leipzig

Brettschneider, Peter

Universitätsbibliothek Trier

Dammeier, Dr. Johanna

Universitätsbibliothek Tübingen

Dubberke, Ina

Hochschulbibliothek am Campus Schöneberg der HWR Berlin

Enders, Carina

Universitätsbibliothek der Technischen Universität Berlin

Finke, Pamela

Universitätsbibliothek der Humboldt-Universität zu Berlin, Zweigbibliothek
Naturwissenschaften

Franik, André

Erzbistum Berlin

Gammert, Jonathan

Universitätsbibliothek Mainz

Genest, Annekathrin

Polnisches Institut Berlin – Bibliothek

Getschmann, Katrin

Glass, Tabea

OsloMet – storbyuniversitetet Universitetesbiblioteket, Oslo

Hemieda, Ranija

Stadtbibliothek Friedrichshain-Kreuzberg, Berlin

Hemme, Felix

Deutsche Zentralbibliothek für Wirtschaftswissenschaften – Leibniz Informationszentrum
Wirtschaft (ZBW), Kiel

Jackenkroll, Melanie

Deutsches Patent- und Markenamt München, Bibliothek

Jurst, Richard

Universitäts- und Landesbibliothek Darmstadt

Karsten, Adienne

Bibliothek des LWL-Archäologie für Westfalen, Münster

Kenter, Jan

Universitäts- und Landesbibliothek Bonn

Kindler, Kirsten

Universität zu Köln, Institut für ausländisches und internationales Strafrecht

Koller, Elisabeth

Stadtbibliothek Friedrichshain-Kreuzberg, Berlin

Krüger, Marion

Universitätsbibliothek Heidelberg

Kuhnert, Dr. Dana

Universitätsbibliothek „Georgius Agricola“ der TU Bergakademie Freiberg

Lemanski, Thorsten

Universitäts- und Landesbibliothek Düsseldorf

Lotz, Jutta

Bibliotheks- und Informationswesen, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR),
Stuttgart

Meinel, Corinna

Universitätsbibliothek Passau

Meixner, Charlotte

Brandenburgische Technische Universität Cottbus-Senftenberg, Universitätsbibliothek

Mikla, Susanne

Universitätsbibliothek Passau

Nüchter, Maria

Württembergische Landesbibliothek Stuttgart

Neumann, Julia

Staatsbibliothek zu Berlin – Preußischer Kulturbesitz

Rösch, Dr. Henriette

Universitätsbibliothek Leipzig

Rechenberger, Heiko

Universitätsbibliothek der Humboldt-Universität zu Berlin

Rupp, Dr. Stephan

Universitätsbibliothek der Universität der Bundeswehr München

Satzinger, Lea

Thüringer Universitäts- und Landesbibliothek Jena

Schmädel, Stephanie

Universitätsbibliothek der Humboldt-Universität zu Berlin

Schmunk, Prof. Dr. Stefan

Hochschule Darmstadt – Fachbereich Media

Schober, Claudia

Universitätsbibliothek Würzburg

Schumann, Natascha

Sebastiani, Valentina

Humboldt-Universität Berlin – Institut für Bibliotheks- und Informationswissenschaft

Stahr, Maik

Uhl, Dr. Susanne

Zentralbibliothek Zürich

Ventzke, Katharina

Vordermayer, Dr. Margaretha

Universitätsbibliothek der Ludwig-Maximilians-Universität München

Weinreich, Meike

Hochschulbibliothek der Fachhochschule Südwestfalen, Soest

Wilke, Christian

Universitätsbibliothek Mannheim

Zielsdorf-Mittag, Sabine

Else-Ury-Bibliothek, Berlin

Veränderungen

Berufliche Veränderungen von VDB-Mitgliedern

Aretz, Beate

bisher: Universitätsbibliothek der Humboldt-Universität zu Berlin

jetzt: Staatsbibliothek zu Berlin – Preußischer Kulturbesitz

Brandtner, Dr. Andreas

bisher: Universitätsbibliothek Mainz

jetzt: Universitätsbibliothek der Freien Universität Berlin

Ganter, Martha

bisher: Stadtbibliothek Neukölln, Helene-Nathan-Bibliothek

jetzt: Universitätsbibliothek der Technischen Universität Berlin

Glitsch, Dr. Silke

bisher: Niedersächsische Staats- und Universitätsbibliothek Göttingen

jetzt: Universitätsbibliothek Kiel

Jungbluth, Anja

bisher: Deutsches Literaturarchiv Marbach, Bibliothek

jetzt: Klassik Stiftung Weimar

Namensänderungen von VDB-Mitgliedern

bisher: Ackermann, Franziska

jetzt: Rapp, Franziska

bisher: Schilling, Susanne

jetzt: Nicolai, Susanne

bisher: Tautz, Barbara

jetzt: Moser, Barbara

Geburtstage

Wir gratulieren sehr herzlich allen Kolleginnen und Kollegen, die einen besonderen Geburtstag begehen!

1. Juli – 30. September 2018

85. Geburtstag

Dr. Laurenz Bösing

Konstanz, ehemals Universitätsbibliothek Trier, am 05.08.2018

Dr. Melitta von Beckerath

Bonn, ehemals Universitäts- und Landesbibliothek Bonn, am 24.08.2018

80. Geburtstag

Dr. Eilhard Cordes

Osnabrück, ehemals Universitätsbibliothek Osnabrück, am 14.07.2018

Dr. Hartmut Harthausen

Dudenhofen, ehemals Landesbibliothekszentrum Rheinland-Pfalz / Pfälzische Landesbibliothek, am 17.07.2018

Dr. Berndt von Egidy

Tübingen, ehemals Universitätsbibliothek Tübingen, am 19.07.2018

Bernd Küppers

Simpelveld/NL, ehemals Universitätsbibliothek der RWTH Aachen, am 30.07.2018

Dr. Heinz-Meinolf Stamm OFM

Paderborn, ehemals Zentralbibliothek des Franziskanerordens Rom, am 26.08.2018

Dr. Ulrich Montag

Ottobrunn, ehemals Bayerische Staatsbibliothek München, am 08.09.2018

75. Geburtstag

Dr. Peter Glanzner

Freiburg, ehemals Bibliothek der Pädagogischen Hochschule Freiburg, am 13.07.2018

Dr. Michael Drucker

Pentling, ehemals Staatliche Bibliothek Regensburg, am 15.07.2018

Reimer Eck

Göttingen, ehemals Niedersächsische Staats- und Universitätsbibliothek Göttingen, am 22.07.2018

Dr. Friedrich Geißelmann

Bad Abbach, ehemals Universitätsbibliothek Regensburg, am 14.08.2018

Wolfgang Binder

Bielefeld, ehemals Universitätsbibliothek Bielefeld, am 06.09.2018

Wolfgang Rhodius

Darmstadt, ehemals Universitäts- und Landesbibliothek Darmstadt, am 09.09.2018

Dr. Gerhart Grüninger

Maria Laach, ehemals Fachinformationszentrum der Bundeswehr, am 10.09.2018

70. Geburtstag

Dr. Berthold Jäger

Fulda, ehemals Bibliothek des Priesterseminars Fulda, am 04.07.2018

Dr. Michael Höppner

Schloß Holte-Stukenbrook, ehemals Universitätsbibliothek Bielefeld, am 05.07.2018

Dr. h.c. Berndt Dugall

Staufenberg, ehemals Universitätsbibliothek Johann Christian Senckenberg Frankfurt, am 22.07.2018

Margrit Lauber-Reymann

München, ehemals Hochschule für den öffentlichen Dienst in Bayern, Fachbereich Archiv- und Bibliothekswesen, am 11.09.2018

Dr. Bernd Lorenz

Lappersdorf, ehemals Hochschule für den öffentlichen Dienst in Bayern, Fachbereich Archiv- und Bibliothekswesen, am 30.09.2018

65. Geburtstag

Dr. Wolf-Dieter Barz

Karlsruhe, Bibliothek des Bundesverfassungsgerichts, am 05.07.2018

Dr. Horst Teichmüller

Universitätsbibliothek der Bauhaus-Universität Weimar, am 21.07.2018

Prof. Dr. Thomas Bürger

Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden, am 24.07.2018

Dr. Ludger Syré

Badische Landesbibliothek Karlsruhe, am 26.07.2018

Dr. Sigrid Hubert-Reichling

Landesbibliothekszentrum Rheinland-Pfalz / Bibliotheca Bipontina, am 10.08.2018

Ewa Dubowik-Baradoy

Tübingen, ehemals Redaktion „Germanistik“, am 03.09.2018

Albert Bilo

Universitätsbibliothek Duisburg-Essen, am 08.09.2018

Henrich Achenbach

Verbundzentrale (VZG) des Gemeinsamen Bibliotheksverbundes Göttingen, am 17.09.2018

60. Geburtstag

Andreas Mälck

Staatsbibliothek zu Berlin – Preußischer Kulturbesitz, am 05.07.2018

Andrea Opitz

Herzog August Bibliothek Wolfenbüttel, am 11.07.2018

Gernot Müller-Kopton

Staats- und Universitätsbibliothek Hamburg Carl von Ossietzky, am 16.07.2018

Brigitte Knödler-Kagoshima

Badische Landesbibliothek Karlsruhe, am 05.08.2018

Dr. Helge Steenweg

Universitätsbibliothek Stuttgart, am 26.08.2018

Dr. Rüdiger Hoyer

Bibliothek des Zentralinstituts für Kunstgeschichte München, am 06.09.2018

Dr. Ulrike Michalowsky

Universitätsbibliothek Potsdam, am 13.09.2018

Prof. Dr. Stefan Gradmann

KU Leuven, Literary Studies Research Unit, am 22.09.2018

Verstorbene VDB-Mitglieder

Wir trauern um unsere Kollegen, die in den letzten Monaten verstorben sind.

Dr. Jürgen Heydrich

* 04.04.1933 † 11.09.2017

Dr. Klaus Weyand

* 08.10.1930 † 09.04.2018

Helmuth Bergmann

* 12.12.1948 † 21.04.2018

Werner Ruddigkeit

* 02.02.1936 † 02.05.2018

Dr. Hannsjörg Kowark

* 03.08.1951 † 14.06.2018

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S255-261>