

Bausteine Forschungsdatenmanagement
Empfehlungen und Erfahrungsberichte für die Praxis von
Forschungsdatenmanagerinnen und -managern

Forschungsdatenmanagement in DFG-Sonderforschungsbereichen – ein Blick in die Praxis

Claudia Engelhardtⁱ

2020

Zitiervorschlag

Engelhardt, Claudia. 2020. Forschungsdatenmanagement in
DFG-Sonderforschungsbereichen – ein Blick in die Praxis. *Bausteine
Forschungsdatenmanagement. Empfehlungen und Erfahrungsberichte für die Praxis von
Forschungsdatenmanagerinnen und -managern* Nr. 1/2020: S. 16-27. DOI:
[10.17192/bfdm.2020.1.8157](https://doi.org/10.17192/bfdm.2020.1.8157).

Dieser Beitrag steht unter einer
[Creative Commons Namensnennung 4.0 International Lizenz \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).

ⁱNiedersächsische Staats- und Universitätsbibliothek Göttingen. ORCID: [0000-0002-3391-7638](https://orcid.org/0000-0002-3391-7638)

1 Abstract

Dieser Artikel stellt die Ergebnisse einer Befragung zum Forschungsdatenmanagement und zur Forschungsdateninfrastruktur in DFG-Sonderforschungsbereichen vor, die im Herbst 2018 im Vorfeld eines Workshops zum selben Thema durchgeführt wurde und an der sich 20 SFBs beteiligten. Die Befragung bestand aus vier offenen Fragen. Diese hatten die Verortung von INF- und anderen Datenmanagementteilprojekten innerhalb der SFBs, ihre Aufgaben, Herausforderungen sowie Policies und Technologien, die eingesetzt werden, zum Thema. Abschließend werden die Ergebnisse mit denjenigen einer ähnlich angelegten Untersuchung aus dem Jahr 2013 verglichen.

2 Einleitung

„Teilprojekte Informationsinfrastruktur“ – oder auch kurz: INF-Projekte – in von der DFG geförderten Sonderforschungsbereichen (SFBs) sollen entsprechend den Vorgaben der DFG „vor allem dem systematischen Management der im Kontext des Sonderforschungsbereichs relevanten Daten“¹ dienen. Dies kann sowohl die Erstellung eines Konzeptes als auch den Aufbau einer entsprechenden Infrastruktur umfassen. Die DFG benennt sechs Ziele, an denen sich die Arbeit in einem INF-Projekt ausrichten kann:

- Zugang zu Forschungsdaten
- Erschließung und Pflege von Forschungsdaten
- Nachnutzung von Forschungsdaten
- Virtuelle Forschungsumgebungen
- Aufbau von interoperablen Komponenten
- Adaption und Implementierung neuartiger Technologien und Verfahren.

Des Weiteren weist sie darauf hin, dass mit den lokalen Infrastruktureinrichtungen, also etwa Bibliotheken, Rechenzentren oder auch Biobanken, kooperiert werden sollte sowie nach Möglichkeit bereits existierende Infrastrukturkomponenten zu verwenden sind.

Wie wird dieser durch die Deutsche Forschungsgemeinschaft gesetzte Rahmen in den einzelnen Sonderforschungsbereichen konkret mit Leben gefüllt und technisch wie organisatorisch umgesetzt? Dieser Frage wurde im Rahmen einer Befragung nachgegangen, die im Vorfeld des Workshops „Forschungsdatenmanagement und -infrastruktur in DFG-Sonderforschungsbereichen“² am 26. und 27. November 2019 in Göt-

¹Vgl. DFG, Merkblatt Sonderforschungsbereiche (DFG-Vordruck 50.06 – 07/18), S. 10, http://www.dfg.de/formulare/50_06/50_06_de.pdf (letzter Zugriff 24.07.2019).

²Die Veranstaltung wurde vom Projekt Göttingen Research Data Exploratory (GRaCE, <http://www.eresearch.uni-goettingen.de/de/content/grace>, letzter Zugriff 24.07.2019) in Kooperation mit Kolleginnen und Kollegen der Universitäten Bielefeld, Dresden, Kiel, Köln und Marburg organisiert. Präsen-

tingen durchgeführt wurde. Dabei wurden nicht nur SFBs mit INF-Projekten berücksichtigt, denn vergleichbare Aktivitäten sind zum Teil auch in anderen Teilprojektformen, etwa Wissenschaftlichen Serviceprojekten (S) oder Zentralen Verwaltungsprojekten (Z), angesiedelt. Diese verschiedenen Formen werden im Folgenden auch unter dem Begriff „Datenmanagement-Teilprojekte“ subsumiert.

Die Erhebung bestand aus vier offenen Fragen:

- Wie ist Ihr Datenmanagement-Teilprojekt innerhalb des SFBs verortet? Welche Aufgaben übernehmen Sie für die anderen Teilprojekte?
- Mit welchen Technologien und Policies unterstützen Sie das Datenmanagement?
- Vor welche Herausforderungen sehen Sie sich in Ihrer praktischen Arbeit gestellt?
- Welche personellen und finanziellen Ressourcen haben Sie für die Bearbeitung Ihrer Aufgaben zur Verfügung?

Die Einladung zur Befragung war Teil der Einladung zum Workshop, die an alle in der GEPRIS-Datenbank³ verzeichneten Sprecherinnen und Sprecher zum damaligen Zeitpunkt laufender SFBs versandt wurde. Es beteiligten sich 20 der 32 beim Workshop vertretenen SFBs⁴, die eine große Bandbreite an Disziplinen abdecken. Unter ihnen sind zehn, denen laut GEPRIS mindestens eine Fachdisziplin der Sozial- und Geisteswissenschaften zugeordnet ist. Sieben SFBs haben einen entsprechenden Bezug zu den Lebenswissenschaften, zwölf zu den Naturwissenschaften und fünf zu den Ingenieurwissenschaften.

Die Antworten wurden mit Hilfe der frei verfügbaren Software OpenCode⁵ einer qualitativen Inhaltsanalyse unterzogen. Die Ergebnisse sind nicht repräsentativ, geben jedoch den Blick auf einen Ausschnitt der aktuellen Praxis des Forschungsdatenmanagements in SFBs frei. Sie werden im Folgenden vorgestellt und abschließend mit den Befunden aus einer ähnlichen, 2013 durchgeführten Untersuchung⁶ verglichen. Neben der fehlenden Repräsentativität ist bei der Interpretation der Ergebnisse zu beachten, dass die Antworten einen unterschiedlichen Detailgrad aufweisen. Dies trifft vor allem

tationen, Poster und weitere Materialien sind unter <https://data.gro.uni-goettingen.de/dataverse/GRAcE> zugänglich (letzter Zugriff 24.07.2019).

³<https://gepris.dfg.de/gepris/> (letzter Zugriff 24.07.2019)

⁴Obwohl die Teilnahme an der Befragung nicht an eine Workshop-Teilnahme gebunden war, erhielten wir nur Antworten von SFBs, die sich auch zur Veranstaltung angemeldet hatten.

⁵<https://www.umu.se/institutionen-for-epidemiologi-och-global-halsa/forskning/open-code/> (letzter Zugriff 24.07.2019).

⁶Diese wurde im Kontext des vom Projekt Radieschen (Rahmenbedingungen einer disziplinübergreifenden Forschungsdateninfrastruktur, <http://www.forschungsdaten.org/index.php/Radieschen>, letzter Zugriff 24.07.2019) veranstalteten „Gemeinsamen Workshops der SFB-INF-Projekte“ durchgeführt. Ein Kurzbericht des Workshops ist verfügbar unter: http://doi.org/10.2312/RADIESCHEN_010. Eine ausführliche Darstellung der Befragungsergebnisse findet sich in: Engelhardt, Claudia, „Forschungsdatenmanagement in DFG-Sonderforschungsbereichen: Teilprojekte Informationsinfrastruktur (INF-Projekte)“, LIBREAS. Library Ideas #23 (2013): 106-130. <https://nbn-resolving.org/urn:nbn:de:kobv:11-100212741>.

auf den Bereich Technologien zu. So kann es sein, dass ein Tool, das von SFB A genannt wurde, in SFB B ebenfalls genutzt wird, auch wenn dieser es in seiner Antwort nicht erwähnt hat. Die Zahlen müssen also nicht unbedingt dem exakten Vorkommen entsprechen, sondern sind eher als Mindestwerte zu betrachten.

3 Verortung der Datenmanagement-Teilprojekte innerhalb des SFBs

Von den 20 antwortenden SFBs verfügen sieben über ein INF-Projekt. In acht Fällen sind entsprechende Aktivitäten in einem Z-Projekt angesiedelt. In jeweils einem weiteren SFB ist das Forschungsdatenmanagement Teil eines S-Projektes bzw. eines wissenschaftlichen Teilprojektes. In Nicht-INF-Teilprojekten sind das Datenmanagement und die Infrastruktur dabei häufig nur einer von mehreren Aufgabenbereichen. Schließlich gibt es unter den befragten SFBs auch drei, in denen relevante Themen und Aufgaben nicht im Rahmen eines spezifischen Teilprojektes bearbeitet werden. In einem dieser SFBs widmet sich eine teilprojektübergreifende Arbeitsgruppe dem Datenmanagement, in den beiden weiteren werden entsprechende Aufgaben von der Geschäftsleitung bzw. der Koordination des SFBs übernommen.

Sechs der an der Befragung teilnehmenden SFBs erwähnen ausdrücklich, dass sie mit SFB-externen Partnern kooperieren. Konkret benannt werden die Bibliothek und das Rechen- bzw. IT-Servicezentrum (beide jeweils zwei Mal) sowie (jeweils einmal) ein universitäres eScience-Zentrum, eine Arbeitsgruppe zum Thema Forschungsdatenmanagement auf Universitätsebene, ein außeruniversitäres Forschungsinstitut und die Partnerhochschulen eines Universitätsverbundes.

4 Aufgaben der Datenmanagement-Teilprojekte

Im Hinblick auf die Aufgaben, die INF- und ähnliche Teilprojekte übernehmen, fallen drei ins Auge, mit denen sich die Mehrheit der 20 an der Befragung teilnehmenden Projekte auf die eine oder andere Art beschäftigt: Dies sind der Aufbau und Betrieb einer Plattform für die zentrale Speicherung und den Austausch der Daten innerhalb des SFBs (18 Mal erwähnt), gefolgt von der Beratung und Unterstützung der Forschenden in Bezug auf Forschungsdatenmanagement und verwandte Themen⁷ (15 Mal genannt) sowie der Durchführung von Workshops und Schulungen⁸ (12 Nennungen). Auf

⁷Die Beratungs- und Unterstützungsangebote waren in fast allen Fällen schon etabliert, nur in zwei der befragten SFBs befanden sie sich zum Zeitpunkt der Befragung noch im Aufbau.

⁸Die Schulungsangebote waren bei neun der befragten SFBs bereits etabliert, in drei Fällen befanden sie sich zum Zeitpunkt der Befragung noch im Aufbau.

die Art der verwendeten Datenplattformen wird im Kapitel „Technologien“ näher eingegangen. In Bezug auf die Beratungs- und Unterstützungsangebote wird in sieben der entsprechenden Antworten nur von Beratung gesprochen, in acht Fällen jedoch ausdrücklich darauf hingewiesen, dass diese im Bedarfsfall auch mit konkreter Unterstützung verbunden ist. Diese muss sich nicht unbedingt nur auf das Datenmanagement allgemein beziehen, sondern kann auch spezifische Fragen angrenzender Themenfelder betreffen, wie beispielsweise methodische Fragen der Datenerhebung und -auswertung, Open Science, Digitalisierung oder Anpassung und Optimierung von Algorithmen.

Neben diesen drei sehr häufig genannten Aktivitäten wurde noch eine Reihe weitere Aufgaben aufgezählt. Hierzu gehören, mit jeweils sechs Nennungen, die Langzeitarchivierung der im SFB erhobenen Daten sowie die Bereitstellung einer Publikationsmöglichkeit, wobei sich Letztere zum Teil auf Textpublikationen, zum Teil aber auch auf die Publikation von Daten und anderen Forschungsartefakten bezieht. Vier der befragten Datenmanagementteilprojekte verfolgen auch eigene Forschungsinteressen. Hiervon bewegen sich zwei in den Digital Humanities und jeweils eines in der Informationstechnologie und im Bereich Informationswissenschaften/Data Science. Ebenfalls vier gaben an, sich mit Datenmodellierung und/oder Metadatenschemata zu beschäftigen. Jeweils drei Mal genannt wurden Bedarfserhebung, Digitalisierung, Entwicklung von Tools sowie Technology Watch in Verbindung mit der Identifikation und Evaluation von für den SFB passenden Infrastrukturkomponenten. Zwei Nennungen sind jeweils für die Entwicklung von (Daten-)Workflows, Datenakquise, Bereitstellung von Datenanalysen und die Betreuung der Webseite zu verzeichnen. Jeweils einmal erwähnt wurden: Entwicklung von Policies, Bereitstellung einer Infrastruktur für Simulationsstudien, Beteiligung bei der Organisation von Veranstaltungen, Gerätebeschaffung sowie Entwicklung eines Lean Process Assessment Models.

5 Policies und Richtlinien

Dreizehn der an der Befragung teilnehmenden Projekte gaben an, dass in ihrem SFB Vorgaben zum Umgang mit Forschungsdaten existieren oder in Vorbereitung sind. In fünf von diesen gab es zum Zeitpunkt der Befragung im Herbst 2018 bereits in Kraft getretene SFB-interne Daten- bzw. Datenmanagement-Policies, in sechs weiteren befanden sich ebensolche in Vorbereitung. In zwei Fällen gibt es keine dezidierte Policy, jedoch sind die Forschenden verpflichtet, ihre Daten im SFB-internen Repository abzulegen. Des Weiteren werden fünf Mal SFB-bezogene Open-Access-Policies und vier Mal hochschulweite Richtlinien erwähnt. Bei Letzteren handelt es sich in jeweils einem Fall um die universitäre Forschungsdatenrichtlinie, die Digitalisierungsstrategie und die Open-Access-Policy; in einem Fall wird sie nicht näher spezifiziert. Schließlich wird auch zwei Mal auf Richtlinien des DFG hingewiesen. In drei Antworten werden Formatvorgaben für Daten thematisiert. Auch Vorgaben in Bezug auf Metadaten werden drei Mal genannt, Empfehlungen für Datenlizenzen einmal.

6 Technologien

Als Plattform für die Speicherung und den Austausch der Forschungsdaten innerhalb des SFBs, deren Aufbau und Betrieb von 18 Datenmanagement-Teilprojekten als Aufgabe genannt wurde, sind verschiedenen Lösungen im Einsatz: Sechs Mal wurden Repository-Systeme (konkret z.B. DSpace, CKAN und DKAN), sieben Mal Datenbanken (konkret z.B. MongoDB, BExIS) genannt. In vier SFBs wird für Datensammlung und -austausch eine Wiki- oder Portalsoftware oder ein Projektmanagementsystem verwendet. Sechs SFBs nutzen selbst entwickelte Daten- bzw. FDM-Plattformen, die zum Teil auch weitere Funktionen, etwa für die Datenveröffentlichung oder die Langzeitarchivierung, beinhalten. In einigen dieser Eigenentwicklungen werden auch ein oder mehrere der zuvor genannten Technologien als Komponenten eingesetzt (und dafür zum Teil angepasst). Es wurden zudem sieben Mal Speichertechnologien benannt, die unabhängig von, zusätzlich zu oder als Ersatz für eine(r) zentrale(n) Datenplattform in Gebrauch sind. Dabei geht es in vier Fällen um Cloud-Storage, zwei Mal um Gruppenlaufwerke bzw. allgemein die universitäre Speicherinfrastruktur sowie einmal um ein Bandarchiv (für sehr große oder sehr sensible Daten).

Mit acht Mal ebenfalls recht häufig angeführt wurden Werkzeuge für das Projektmanagement und die projektinterne Kollaboration. Zudem werden Tools für das Code Management und die Versionskontrolle wie beispielsweise Git und Subversion (fünf Nennungen), Werkzeuge zur Datenanalyse (drei Nennungen) und WebGIS-Tools (drei Nennungen) angewendet. Ein Werkzeug, das die Datenmanagementplanung unterstützt wurde lediglich einmal erwähnt. Außerdem wurde im entsprechenden SFB zum Zeitpunkt der Befragung dessen Einsatz erst erwogen; eine aktive Nutzung fand noch nicht statt. In sechs Antworten wird auf die Verwendung von Persistenten Identifikatoren (PIDs) hingewiesen, wobei in allen Fällen DOIs für die Referenzierung von Daten und in drei Fällen zusätzlich ORCIDs für Forschende eingesetzt werden.

Darüber hinaus wurde in vier Antworten betont, dass auf die Nutzung von Open-Source-Software Wert gelegt wird. Ebenfalls vier SFBs berichten von der Nachnutzung von Infrastrukturkomponenten. Dies betrifft in einem Fall bestehende Services des Rechenzentrums, in einem weiteren ein disziplinspezifisches Werkzeug zur Textanalyse. In zwei Fällen wird Infrastruktur aus einem anderen SFB nachgenutzt. In diesem Zusammenhang ist auch ein Beispiel erwähnenswert, in dem gute Erfahrungen in einem INF-Projekt ein zusätzliches, von der Universität finanziertes FDM-Pilotprojekt inspiriert haben, aus dem nach dem Projektende ein universitäres FDM-Kompetenzzentrum hervorging, das personell mittlerweile mit über fünf unbefristet besetzten Vollzeitäquivalenten ausgestattet ist.

7 Herausforderungen

Als eine der größten und häufigsten Herausforderung werden Probleme hinsichtlich der Akzeptanz des Datenmanagements und der Arbeit der entsprechenden Teilprojekte oder Arbeitsgruppen geschildert. Insgesamt wird in 15 Antworten darauf verwiesen. Dabei geht es in den meisten (konkret in zwölf) Fällen darum, dass auf Seiten der Forschenden das Bewusstsein für die Notwendigkeit und den Mehrwert des Forschungsdatenmanagements und mithin für die Arbeit entsprechender Teilprojekte eher gering ausgeprägt ist. Bewusstseinsfördernde Maßnahmen stellen daher für viele dieser Teilprojekte einen wichtigen Teil ihrer Arbeit dar. Als in dieser Hinsicht förderlich wurden die regelmäßige Teilnahme an Projekttreffen, Sitzungen und ähnlichen Veranstaltungen, das regelmäßige Vorstellen der Arbeit und Ergebnisse des Datenmanagement-Teilprojekts, das Veranstalten von Workshops zum Thema FDM sowie ein enger Austausch und eine gute Vernetzung mit den wissenschaftlichen Teilprojekten und der SFB-Leitung beschrieben.⁹ Selbst wenn eine gewisse Sensibilisierung erreicht wurde, kann es dennoch vorkommen, dass die Angebote der Datenmanagement-Teilprojekte nur zögerlich genutzt werden. Ein Grund hierfür ist auch, dass ein sorgfältiges Datenmanagement einen gewissen Aufwand auf Seiten der Forschenden erfordert, was in vier Antworten als Hemmnis angeführt wurde. Auch unterschiedliche Wissenschaftskulturen sowie etablierte Forschungsprozesse, die durch hinzukommendes Datenmanagement geändert werden müssten, spielen zum Teil eine Rolle und wurden zwei Mal erwähnt. Ein Datenmanagement-Teilprojekt, das selbst auch Forschung betreibt, berichtete zudem, dass es schwierig sei, von den anderen Teilprojekten als Forschungspartner akzeptiert zu werden.

Zehn Mal wurden verschiedene Facetten von Heterogenität als Herausforderung genannt. Dies bezieht sich vor allem auf die zahlreichen unterschiedlichen Disziplinen, Forschungsmethoden und -prozesse sowie Datenarten, die in einem SFB vereint sind (acht Nennungen). Es gestaltet sich schwierig, einheitliche – technische wie informationswissenschaftliche – Lösungen zu entwickeln, die den vielfältigen daraus resultierenden Anforderungen gerecht werden. Außerdem wurde in diesem Zusammenhang auch drei Mal auf das SFB-intern sehr unterschiedlichen Niveau der Kenntnisse und Kompetenzen in Bezug auf Forschungsdaten hingewiesen.

Eine weitere Gruppe von Herausforderungen, die insgesamt acht Mal erwähnt wurde, lässt sich unter dem Begriff Ressourcenmangel zusammenfassen. Dies bezieht sich in erster Linie auf Personalmittel, zum Teil auch auf Sachmittel. In vier Antworten wird als ein Grund für diesen Mangel angegeben, dass an den entsprechenden Universitäten noch keine ausreichende FDM-Infrastruktur (die INF-Projekte ja angehalten sind zu nutzen) sowie -Beratung vorhanden ist und dies vom SFB-Datenmanagement-Teilprojekt kompensiert werden muss, was dessen Ressourcen jedoch übersteigt. Ein

⁹Zudem sei es vorteilhaft, wenn in den einzelnen Teilprojekten „verbündete“ Forschende gewonnen werden könnten, die beispielsweise neue vom INF-Projekt bereitgestellte Infrastrukturkomponenten testen, was sich im besten Fall auch positiv auf das FDM-Bewusstsein der anderen Wissenschaftlerinnen und Wissenschaftler im Teilprojekt auswirkt.

Datenmanagement-Teilprojekt, das neben FDM auch andere Aufgaben für den SFB übernimmt, sah in diesem Umstand ebenfalls einen Grund für nicht ausreichende Personalkapazität.

Jeweils sieben Mal wurden einerseits technische, andererseits informationswissenschaftliche Herausforderungen geschildert. Technische Herausforderungen resultieren zum einen daraus, geeignete Lösungen für die oben bereits angedeuteten heterogenen Herausforderungen eines SFBs zu finden. Zum anderen ist auch die konkrete Implementierung solcher Lösungen nicht immer einfach. Die informationswissenschaftlichen Herausforderungen sind divers. Zum Teil sind auch sie in der Heterogenität der Anforderungen begründet, die beispielsweise die Entwicklung von einheitlichen Workflows oder Metadatenschemata erschwert. Auch die Definition der Aufgaben des Datenmanagements sowie die Auswahl von Forschungsdaten für die Archivierung oder Veröffentlichung wurden in diesem Kontext genannt. Zudem fallen in diese Kategorie (zu Projektbeginn) fehlende FDM-Expertise im Datenmanagement-Teilprojekt selbst sowie fehlende Expertise zur Beratung von Forschenden, z.B. bezüglich der Nutzung von Software.

Mit rechtlichen Aspekten verbundene Herausforderungen wurden fünf Mal erwähnt. Hier spielen datenschutzrechtliche Belange, insbesondere im Zusammenhang mit personenbezogenen Daten, eine Rolle – ebenso wie lizenzrechtliche Hürden bei der Nachnutzung und Verfügbarmachung von Daten, etwa von Textkorpora oder Digitalisaten. Schwierigkeiten, geeignetes Personal für Stellen im Bereich Datenmanagement zu finden, wurden drei Mal angeführt. Weiterhin wurde zwei Mal auf Schwierigkeiten in Bezug auf die Kooperation verschiedener beteiligter Parteien verwiesen, wobei in beiden Fällen unter anderem auch der Umgang mit den oft divergierenden Interessen und Anforderungen, die von unterschiedlicher Seite an das Datenmanagement-Teilprojekt herangetragen werden, als herausfordernd beschrieben wurde. Vier weitere Herausforderungen¹⁰ wurden jeweils einmal genannt.

8 Personelle und finanzielle Ressourcen

Hinsichtlich der personellen Ausstattung mit Stellen für das Datenmanagement und Infrastruktur lässt sich eine Tendenz mit Abweichungen nach unten und oben ausmachen. Der häufigste Fall, der auf acht der an der Befragung teilnehmenden SFBs zutrifft, ist eine volle Stelle, zusätzlich zum Teil Mittel für eine oder zwei studentische oder wissenschaftliche Hilfskräfte. In sieben Fällen handelte es dabei sich um eine Postdoc-, in einem um eine IT-Stelle. Zusätzliche Hilfskraftstellen wurden in sechs Mal angegeben – dabei geht es drei Mal um eine und drei Mal um zwei Hilfskraftstellen.

¹⁰Diese sind: 1) die Größe des und die Entfernung zum Untersuchungsgebiet, 2) sehr große Datenmengen (im dreistelligen Terabyte-Bereich, 3) Start des Datenmanagement-Teilprojekts erst in der zweiten Förderphase, wodurch die Datenmodellierung nicht von Beginn an erfolgen konnte, 4) die Anbindung an das WDC wird unter Sachmitteln, statt unter Publikationskosten veranschlagt.

Eine weitere Kombination, die drei Mal erwähnt wurde, ist eine volle Stelle für die wissenschaftliche Mitarbeit und eine volle (zwei Nennungen) oder halbe (eine Nennung) IT-Stelle. In zwei Fällen stehen insgesamt mehr als zwei Vollzeitäquivalente (VZÄ) zur Verfügung, in fünf SFBs weniger als ein VZÄ (Hilfskraftmittel nicht mitgerechnet). Von Letzteren haben jedoch drei kein eigenes Datenmanagement-Teilprojekt. In drei Fällen konnte aus den Angaben der genaue Stellenumfang nicht ermittelt werden. Neben den oben genannten Postdoc-, IT- und Hilfskraftstellen wurden zudem in einigen Fällen Eigenmittel, etwa in Form von PI-Stellenanteilen, Qualifikationsstellen und (SFB-interne) Projektstellen erwähnt.

Angaben zu Sachmitteln wurden nur von neun Befragten gemacht. Am häufigsten wurden Mittel für Speicher oder andere Infrastruktur (fünf Nennungen) sowie Reisemittel (vier Nennungen) erwähnt. Darüber hinaus wurden jeweils einmal Mittel für Workshops, für Fortbildungen sowie die Anbindung an das World Data System (WDS) angegeben. In zwei Fällen stehen dem Datenmanagement-Teilprojekt keine eigenen Sachmittel zur Verfügung bzw. wurde auf zentrale Mittel des SFB verwiesen.

Schließlich wird in einem SFB während der Laufzeit von der Universität Ausstattung in Form von Speicherkapazität im zweistelligen Terabyte-Bereich zur Verfügung gestellt.

9 Vergleich 2013 – 2018

Lassen sich Gemeinsamkeiten oder Unterschiede zwischen den Befunden aus den Jahren 2013¹¹ und 2018 feststellen? Aus Platzgründen wird der Vergleich hauptsächlich in tabellarischer Form durchgeführt und auf eine ausführliche Erläuterung in Fließtext verzichtet. Neben den im Folgenden angeführten möglichen Erklärungen für Unterschiede sollte ebenfalls bedacht werden, dass es sich nicht um repräsentative Erhebungen handelt und manche Differenzen auch einfach dem Zufall und/oder dem unterschiedlichen Detailgrad der Antworten geschuldet sein könnten.

Bei Betrachtung der Aufgaben (Tabelle 1) zeigen sich viele Parallelen sowie einige kleinere Unterschiede. Die am häufigsten genannte Aufgabe war zu beiden Erhebungszeitpunkten der Aufbau einer zentralen Plattform für die Speicherung und den Austausch der Forschungsdaten des SFBs. Auffällig ist, dass die Beratung und Unterstützung der Wissenschaftlerinnen und Wissenschaftler in FDM-Fragen sowie das Durchführen von Workshops und Schulungen in den Aufgabenbeschreibungen aus dem Jahr 2018 erheblich an Stellenwert gewonnen hat. Ein weiterer Unterschied, der hervorsteicht, ist die Bereitstellung von kollaborativen Arbeitsumgebungen, die 2013 zehn Mal angeführt wurde, 2018 aber in den Aufgaben gar nicht mehr vorkommt. Eine

¹¹Die Ergebnisse der im Jahr 2013 durchgeführten Befragung finden sich in ausführlicher Form in: Engelhardt, „Forschungsdatenmanagement in DFG-Sonderforschungsbereichen, 106-130.

mögliche (teilweise) Erklärung hierfür könnte sein, dass kollaborative Elemente zunehmend als Teil der Datenspeicher- und Austauschplattform verstanden bzw. implementiert werden. Wirft man einen Blick auf die eingesetzten Technologien (Tabelle 3), wurden Werkzeuge zur Unterstützung des Projektmanagements und des kollaborativen Arbeitens auch 2018 häufig genannt.

2013	n=18	n=20	2018
Datenspeicher-/Austauschplattform	13	18	Datenspeicher-/Austauschplattform
Kollaborative Arbeitsumgebung	10	15	Beratung und Unterstützung
Beratung, Unterstützung, Schulung	7	12	Workshops, Schulungen
Entwicklung von Tools	7	6	Langzeitarchivierung
Publikation(smöglichkeit)	6	6	Publikation(smöglichkeit)
Langzeitarchivierung	4	4	Forschung
IT-Administration	3	4	Metadaten/Datenmodellierung
Entwicklung von Formaten	2	3	Identifikation
			Infrastrukturkomponenten
Forschung	2	3	Bedarfserhebung
Datenanalyse	2	3	Digitalisierung
Digitalisierung	1	3	Entwicklung von Tools
		2	Entwicklung von (Daten-)Workflows
		2	Datenakquise
		2	Bereitstellung von Datenanalysen
		2	Betreuung der Webseite
		1	Sonstiges

Tabelle 1: Aufgaben

Ein ähnlicher Mix von Gemeinsamkeiten und Unterschieden zeigt sich bei den Policies und Richtlinien (Tabelle 2). Führend sind in beiden Jahren Daten(management)-Policies bzw. -Richtlinien. Der Stellenwert von Richtlinien zu Open Source und Metadaten scheint 2018 etwas niedriger zu sein als noch 2013. Richtlinien zum Datenschutz, 2013 drei Mal erwähnt, kommen 2018 in den Antworten nicht vor. Möglicherweise werden diese Elemente zunehmend in die allgemeine Datenmanagement-Policy integriert. Im Jahr 2018 neu hinzugekommen sind Open-Access-, Hochschul- und Förderer-Policies.

2013	n=18	n=20	2018
Daten-/Datenmanagement-Policy	8	11	Daten-/Datenmanagement-Policy
Metadatenvorgaben	7	5	Open-Access-Policy
Open Source	7	4	Open Source
Richtlinien Datenschutz/-sicherheit	3	4	Hochschul-Policy
Terms of Use-Vorlagen	3	3	Formatvorgaben
Datenablagepflicht im Repository	2	3	Metadatenvorgaben
Orientierung an Nutzenden	2	2	Datenablagepflicht im Repository
		2	Förderer-Policy
		1	Lizenzvorgaben/-empfehlungen

Tabelle 2: Policies und Richtlinien

In Bezug auf die verwendeten Technologien (Tabelle 3) überwiegen eindeutig die Gemeinsamkeiten. Lediglich Werkzeuge für das Code Management und WebGIS-Tools waren 2018 als neue Kategorien zu verzeichnen.

2013	n=18	n=20	2018
Projektmanagement- /Kollaborationstools	8	8	Projektmanagement- /Kollaborationstools
Speichertechnologien	6	7	Datenbanken
Eigenentwicklungen	5	6	Repository-Systeme
Datenbanken	5	6	Eigenentwicklungen
Repository-Systeme	4	6	Speichertechnologien
Persistent Identifier	1	6	Persistent Identifier
Tools zur Datenanalyse	1	5	Tools f. Code Management- /Versionskontrolle
Sonstige	4	3	Tools zur Datenanalyse
		3	WebGIS-Tools
		1	Sonstige

Tabelle 3: Technologien

Die Befunde zu den Herausforderungen (Tabelle 4) gleichen sich ebenfalls sehr stark. Der Punkt „Metadaten“, der 2013 noch als eigener Punkt angeführt ist, ist in den Befunden von 2018 in den etwas weiter gefassten „informationswissenschaftlichen Herausforderungen“ enthalten.

2013	n=18	n=20	2018
Heterogenität	11	15	Akzeptanz
Akzeptanz	11	10	Heterogenität
Technische Herausforderungen	6	8	Ressourcenmangel
Ressourcenmangel	5	7	Technische Herausforderungen
Metadaten	4	7	Informationswissenschaftliche Herausforderungen
(Un-)Gleichzeitigkeit ¹²	3	5	Rechtliche Aspekte
Rechtliche Aspekte	3	3	Schwierigkeiten bei Personalakquise
Schwierigkeiten bei Personalakquise	1	2	Schwierige Kooperation
		4	Sonstiges

Tabelle 4: Herausforderungen

10 Fazit

Insgesamt überwiegen beim Vergleich zwischen 2013 und 2018 die Parallelen, insbesondere beim Blick auf die meistgenannten Kategorien. Bezüglich der Aufgaben der Datenmanagement-Teilprojekte stechen hier der Aufbau und Betrieb einer Plattform für Datenspeicherung und -austausch sowie die Beratung, Unterstützung und Schulung der Forschenden des SFBs in Bezug auf Datenmanagementfragen hervor. Im Bereich Policies wurden sowohl 2013 als auch 2018 am häufigsten Daten- bzw. Datenmanagement-Policies genannt. Es sind eine Reihe von Technologien im Einsatz. In beiden untersuchten Jahren wird die Liste von Projektmanagement- und Kollaborationstools angeführt, jedoch dicht gefolgt von einer Gruppe weiterer Komponenten (Repository-Systeme, Datenbanken, Speichertechnologien, Eigenentwicklungen), die sich alle ungefähr auf demselben Häufigkeitslevel bewegen. Als die größten Herausforderungen wurden in beiden Jahren übereinstimmend das Problem der (mangelnden) Akzeptanz des Datenmanagements und der Arbeit der Datenmanagement-Teilprojekte durch die Forschenden sowie der Umgang mit Heterogenität (an Disziplinen, Methoden, Datenarten etc. innerhalb des SFBs) beschrieben.

¹²Dies bezieht sich auf das Problem, dass die Datenmanagement-Projekte zeitgleich mit den anderen Teilprojekten starten. In der Regel braucht es eine gewisse Zeit zur Konzeption und zum Aufbau der Infrastruktur und anderer Angebote, welche jedoch von den Forschenden von Beginn an benötigt werden.