



Aufbau und Erprobung von Bausteinen für ein effektives
und effizientes Forschungsdatenmanagement

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

PROJEKTBERICHT 1.1:

Forschungsdatenmanagement an der Universität Jena: Interviews zum Stand und Bedarf bei Verbundprojekten

Bettina Hesse, Markus Baaske, Roman Gerlach, Birgitta König-Ries

Zitiervorschlag:

Hesse, Bettina; Baaske, Markus; Gerlach, Roman; König-Ries, Birgitta (2017):
Forschungsdatenmanagement an der Universität Jena: Interviews zum Stand und Bedarf bei
Verbundprojekten. Bericht 1.1. eeFDM-Projekt (BMBF), Jena.

Schlagwörter:

Forschungsdaten, Forschungsdatenmanagement (FDM), Universität, Befragung, Interview,
Verbundprojekt, Status, Bedarf, zentrale Serviceleistungen

Autorenanschrift:

Kontaktstelle Forschungsdatenmanagement
am Michael-Stifel-Zentrum Jena
Friedrich-Schiller-Universität Jena
Ernst-Abbe-Platz 2
07743 Jena

Das Projekt eeFDM „Aufbau und Erprobung von Bausteinen für ein effektives und effizientes
Forschungsdatenmanagement“ wird durch das Bundesministerium für Bildung und Forschung
(BMBF) im Rahmen der Förderrichtlinie zur Erforschung des Managements von Forschungsdaten in
ihrem Lebenszyklus an Hochschulen und außeruniversitären Forschungseinrichtungen gefördert
(Förderkennzeichen 16FDM009). Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei
den Autoren.

November 2017

Version 1.1

Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung - Weitergabe unter
gleichen Bedingungen 4.0 International Lizenz



Zusammenfassung

Die Friedrich-Schiller-Universität Jena (FSU) ist die größte Hochschule in Thüringen. Sie ist bestrebt, ihre Angehörigen mit geeigneten Angeboten zum Forschungsdatenmanagement (FDM) entlang des gesamten Datenlebenszyklus zu unterstützen und damit eine Vorbildfunktion im Bundesland zu übernehmen. In den letzten Jahren hat die FSU begonnen, dafür die organisatorischen Voraussetzungen zu schaffen. Ziel ist nun die Erhöhung von Effektivität und Effizienz des FDM.

Dafür wurden in einem ersten Schritt 10 Verbundprojekte in teilstrukturierten Interviews zu Ihrem Stand im FDM befragt. Dabei wurde darauf geachtet, dass ein breites Spektrum an Disziplinen, Laufzeitabschnitten und FDM-Vorerfahrung abgedeckt wurde. Es sollte eruiert werden, welche Strukturen im FDM bereits vorhanden sind und an welcher Stelle die FSU Unterstützung leisten kann, welche Dienste lokal und fachspezifisch sein sollten und wo eine Zentralisierung und Generalisierung sinnvoll ist. Es wurden sowohl die technische Infrastruktur als auch Angebote für Beratung und Schulungen betrachtet.

Die Datenquellen in den befragten Projekten sind vielfältig. Das geschätzte Datenaufkommen zeigt ein breites Spektrum (Megabyte- bis Petabyte-Bereich) ebenso wie die verwendeten Datenformate (allgemeine Standards bis herstellereigene). Im Allgemeinen ist FDM als Thema präsent und in den meisten Projekten gibt es eine Data Policy oder ein Äquivalent. Auch in der Graduiertenausbildung ist das Thema angekommen. Die Bedeutung von Qualitätssicherung und guten Metadaten ist erkannt. Insbesondere die „10-Jahres Regel“ zur Datenarchivierung aus den DFG Empfehlungen zur Sicherung guter wissenschaftlicher Praxis (1998, 2013) wurden regelmäßig genannt. Die 2015 von der DFG und 2017 von der Universität Jena herausgegebenen Leitlinien zum Umgang mit Forschungsdaten und Maßnahmen für die darin geforderte Nachnutzbarkeit von Forschungsdaten sind dagegen noch nicht ausreichend verankert. Hier besteht ein gewisser Aufklärungsbedarf, um diese Leitlinien in der Praxis hinreichend zu etablieren.

Verbundprojekte, insbesondere mit Beteiligung außeruniversitärer Forschungseinrichtungen, decken den Bedarf an IT-Infrastruktur mehrheitlich aus den gut ausgestatteten Ressourcen der einzelnen Partner. Nur wenige Verbundprojekte implementieren zentrale projektspezifische Dateninfrastrukturen für die Projektlaufzeit. Von besonders großem Interesse ist das Thema Langzeitspeicherung. Auf diesem Gebiet konnten offensichtlich vielfältige Fragen der (praktischen) Umsetzbarkeit noch nicht abschließend geklärt werden. Zu beobachten war allerdings, dass durch die sogenannten INF-Projekte tendenziell die Arbeitsabläufe, Verantwortlichkeiten und Kompetenzen eines verantwortungsvollen Umgangs mit Forschungsdaten strukturierter gebündelt werden konnten. Weiterhin besteht ein erhöhter Bedarf an praxisorientierten, „guten“ Best Practice Beispielen sowie in individueller, technischer und rechtlicher Beratung zum FDM. Gefordert wurde hierbei, dass der Aufwand durch ein umfängliches FDM stets in einem vertretbaren Verhältnis zu dessen Mehrwert liegen sollte. Hier spielte insbesondere die Verstärkung der Angebote für ein verantwortungsvolles FDM durch spezialisiertes und langfristig verfügbares Personal eine wesentliche Rolle.

Summary

The Friedrich Schiller University Jena (FSU) is the largest university in the state of Thuringia. It aims to support its members with adequate services for research data management (RDM) over the entire data life cycle. Thereby, it intends to act as a model in the state. In the last years, the FSU started to establish the administrative background. Now, the aim is to make RDM more effective and efficient.

In a first step, the status quo in RDM was investigated in 10 collaborative projects via partially structured interviews. The projects were selected to cover a wide range of disciplines, time points in the life span and RDM experience. The questions to answer were: Which RDM structures are already implemented and where is support needed? Which services should be provided locally and discipline-specific and whether there is potential for generic central services? The focus thereby was on technical infrastructure, support and training.

The sources of research data in the projects under survey are manifold. The estimated total data volume varies from megabytes to petabytes. The same applies to data formats (general, vendor specific or even self-made standards). Generally speaking, there is a certain awareness of RDM practices and most of the projects do have a data policy or something equivalent in place. To some degree, RDM has become part of the education of graduate students. The significance of quality assurance and appropriate metadata is now recognised. In particular, the 'rule of 10 years' concerning data archiving of the DFG Guideline on Safeguarding Good Scientific Practice (DFG, 1998 & 2013) has been mentioned frequently. In contrast, the Guidelines on the Handling of Research Data, published by the DFG (2015) as well as the University of Jena (2017), and the demand for reusability stated herein have not been implemented sufficiently until now. There is still a need for information and clarification in order to put this guideline into practice.

Collaborative projects, especially if non-university research facilities are involved, mostly use the IT-infrastructure available in the participating institutions, which are typically well equipped. Only a few collaborative projects implement a dedicated central data infrastructure for the project life time. On the contrary, providing long-term storage facilities for data was of great interest. It seems there are a number of open questions in this regard, especially concerning the practical implementation. However, the so-called "INF projects" (DFG CRC) seem to encourage responsible conduct of research data management in terms of improved organisation, structured workflows, and well defined responsibilities. In addition, there is a demand for practical training courses, best practice examples as well as individual, technical and legal advice. It was pointed out, that the effort required to conduct comprehensive RDM should always be proportional to its benefits and objectives. In this context, RDM services ought to be institutionalized and permanently available, same as well trained permanent staff.

Inhalt

Inhalt	5
Abkürzungen	6
1 Einleitung.....	7
1.1 Hintergrund	7
1.2 Motivation.....	7
1.3 Befragungen zum FDM an anderen deutschen Hochschulen	8
1.4 Ziel	9
2 Methoden.....	10
2.1 Projektauswahl.....	10
2.2 Interviewleitfaden	12
2.3 Durchführung der Befragung	12
3 Ergebnisse.....	12
3.1 Aktueller Stand zum Forschungsdatenmanagement	12
3.2 Charakterisierung der Daten	13
3.3 IT-Infrastruktur und Service	14
3.4 Langzeitarchivierung, Nachnutzbarkeit und Publikation von Daten.....	15
3.5 Lehr- und Kursangebote	16
3.6 Anmerkungen und Wünsche.....	17
4 Diskussion.....	18
Anmerkungen zur Interpretation	18
Stand und Bedarf in den Verbundprojekten an der FSU.....	19
5 Ausblick.....	20
Danksagung	22
Referenzen	23
Anhang	25
Interviewleitfaden	25

Abkürzungen

AG	Arbeitsgruppe
BEXIS	Biodiversitäts-Exploratorien Informationssystem
BMBF	Bundesministerium für Bildung und Forschung
DFG	Deutsche Forschungsgemeinschaft
EBI	European Bioinformatics Institute
EOSC	European Open Science Cloud
FDM	Forschungsdatenmanagement
FLI	Leibniz-Institut für Alternsforschung - Fritz-Lipmann-Institut
FOR	Forschergruppe
FSU	Friedrich-Schiller-Universität Jena
GEO	Gene Expression Omnibus
GFBio	German Federation for Biological Data
INF-Projekt	Programmelement Informationsmanagement und -infrastruktur in SFBs
IT	Informationstechnik
KS	Kontaktstelle
LIMS	Laborinformations- und Management-System
LIS	Wissenschaftliche Literaturversorgungs- und Informationssysteme
MPI	Max-Planck-Institut
MSCJ	Michael Stifel Center Jena
NCBI	National Center for Biotechnology Information
NFDI	Nationale Forschungsdateninfrastruktur
RFII	Rat für Informationsinfrastrukturen
SFB	Sonderforschungsbereich
SFT	Servicezentrum Forschung und Transfer
ThULB	Thüringer Universitäts- und Landesbibliothek
UKJ	Universitätsklinikum Jena
URZ	Universitätsrechenzentrum
VPF	Vizepräsidium Forschung
ZKS	Zentrum für Klinische Studien

1 Einleitung

1.1 Hintergrund

Die Friedrich-Schiller-Universität Jena (FSU) ist eine multidisziplinäre Hochschule (10 Fakultäten, s. auch Methoden: Projektauswahl) mit einer ausgeprägten Kultur der Inter- und Transdisziplinarität. Die Kooperation mit außeruniversitären Forschungseinrichtungen, der regionalen Wirtschaft sowie hochrangigen Kulturinstitutionen umfasst ein breites Spektrum aktueller Forschungsthemen. Dem wird mit drei Profillinien strukturell Ausdruck verliehen: „Light“ (Optik und Photonik, Innovative Materialien, Energiespeicher) – „Life“ (Mikrobiologie und Infektionsforschung, Biodiversität und Bio-Geo-Interaktionen, Altersforschung) – „Liberty“ (Sozialer Wandel, Aufklärung, Romantik, Osteuropa, Zeitgeschichte).

Die FSU ist die größte Universität des Landes Thüringen und diejenige unter den Thüringer Hochschulen, die bisher die umfangreichsten Erfahrungen im Bereich Forschungsdatenmanagement (FDM) gesammelt hat. Bereits jetzt findet vielfach eine Abstimmung mit anderen Akteuren am Standort Jena (insbesondere den außeruniversitären Forschungseinrichtungen, etwa im Rahmen gemeinsamer Verbundprojekte) und in Thüringen (etwa über Verbände von Rechenzentren und Bibliotheken) statt. Diese Vernetzung soll auf allen Ebenen (strategisch, organisatorisch sowie operationell) weiter ausgebaut werden. Die FSU strebt hier eine Vorreiterrolle im Bundesland an.

In den letzten Jahren hat die FSU begonnen, dafür die organisatorischen Voraussetzungen zu schaffen: Das Thema FDM ist auf Ebene der Hochschulleitung verankert und fällt in die Zuständigkeit des Vizepräsidiums Forschung (VPF). Strategische Vorgaben werden von einer vom Forschungsausschuss des Senats der FSU eingesetzten Arbeitsgruppe (AG FDM) erarbeitet, in den universitären Gremien vertreten und in ihrer Einhaltung überwacht. Die AG FDM ist aus Vertreterinnen und Vertretern der Wissenschaft, des VPF, der zentralen Serviceeinrichtungen (Universitätsrechenzentrum (URZ), Thüringer Universitäts- und Landesbibliothek (ThULB), Servicezentrum Forschung und Transfer (SFT), Rechtsamt, Zentrum für Klinische Studien (ZKS)) sowie der im März 2015 gegründeten Kontaktstelle Forschungsdatenmanagement (KS FDM) zusammengesetzt. Die KS FDM wurde jüngst an das ebenfalls 2015 gegründete Michael Stifel Center Jena (MSCJ) angegliedert, das die Förderung datengetriebener und simulationsgestützter Forschung zum Ziel hat.

1.2 Motivation

Die FSU misst Forschungsdaten als wesentlicher Grundlage wissenschaftlicher Erkenntnis eine hohe Bedeutung bei und ist bestrebt, ihre Angehörigen mit geeigneten Angeboten zum FDM entlang des gesamten Lebenszyklus von Forschungsdaten zu unterstützen. Die Forschenden sollen von Routineaufgaben der Datenhaltung entlastet werden, um sich disziplinspezifischer Datenintegration und -analyse, also ihren „eigentlichen“ Aufgaben widmen zu können.

Im Rahmen der Strategieentwicklung fand zunächst eine Statusanalyse durch die AG FDM statt. Es wurde eine Gegenüberstellung an der FSU existierender Angebote und des gewünschten Soll-Zustands erarbeitet. Berücksichtigt wurden dabei sowohl Verbundprojekte, als auch Einzelvorhaben. Daraufhin wurden eine Leitlinie sowie eine Handlungsempfehlung zum FDM entwickelt, die am 20. Dezember 2016 vom Senat verabschiedet wurde. Die nächste Aufgabe der AG FDM wird es sein, eine abschließende Empfehlung zur konkreten Organisationsstruktur zu entwickeln, die neben der KS FDM alle relevanten Serviceeinrichtungen einbezieht und eine enge Kopplung an die Wissenschaft und die Bedürfnisse der Forschenden garantiert.

Der Ausbau von Infrastrukturen wurde punktuell bereits begonnen. URZ und ThULB stellen Angebote zur Sicherung und Verfügbarmachung von Forschungsdaten nach Abschluss der Arbeiten bereit. Dieses institutionelle Repositorium soll als Fallback-Lösung für die Fälle dienen, in denen keine geeigneten fachspezifischen Repositorien im nationalen oder internationalen Umfeld existieren. Außerdem genießt die FSU bereits heute nationale Sichtbarkeit bei Entwicklung und Betrieb von FDM-Plattformen für die Biodiversitätsforschung (Profillinie „Life“) und ist an GFBio, einem DFG-LIS geförderten Projekt, das den Aufbau einer nationalen Infrastruktur zur Langzeitsicherung biologischer Daten zum Ziel hat, beteiligt. Die FSU strebt eine aktive Rolle beim Aufbau einer nationalen Forschungsdateninfrastruktur insbesondere im Bereich der Biodiversitätsforschung an.

Der Aufbau von Services zu sämtlichen Aspekten dieses Lebenszyklus folgt an der FSU dem Ansatz, lokal vorhandene Dienste anzupassen, nationale oder internationale Angebote einzubinden, und Eigenentwicklungen nur dann vorzunehmen, wenn tatsächlich keine Angebote bestehen oder fachspezifische Besonderheiten dies erfordern. Die FSU sieht sich hier als aktiver Partner in einem kollaborativen Umfeld mit anderen Forschungseinrichtungen und beteiligt sich über Projekte aktiv an den allgemeinen Anstrengungen zum Aufbau von Services und Infrastruktur im Umgang mit Forschungsdaten.

Strategisch strebt die FSU die Erhöhung von Effektivität und Effizienz des FDM an. Erhöhte Effektivität bedeutet: besseres FDM durch vollständige Abdeckung des Lebenszyklus, das in der Folge u.a. zu verbesserter Reproduzierbarkeit von Forschungsergebnissen, gesteigerter Nachnutzbarkeit von Daten, Erhöhung des Impacts der Forschung an der FSU und erhöhter Drittmittelfähigkeit führt. Erhöhte Effizienz bedeutet: „billigeres“ FDM, das sich u.a. durch Ressourcenschonung durch Synergieeffekte beim Datenmanagement, vermiedene Mehrfacherhebung von Daten, gemeinschaftliche Nutzung von Ressourcen, Einbindung in nationale und internationale Infrastrukturen und verminderte zeitliche Belastung von Forschenden auszeichnet. Dazu gehört auch die Weiterentwicklung von Kompetenzen über unterschiedliche Beratungs- und Schulungsangebote, die durch die KS FDM bereits begonnen wurde und weiter intensiviert werden soll.

1.3 Befragungen zum FDM an anderen deutschen Hochschulen

In den letzten vier Jahren wurden an einigen deutschen Hochschulen Status- und Bedarfserhebungen zum Thema FDM durchgeführt. Verantwortlich zeichneten sich hier sehr häufig die mit den Hochschulen assoziierten Bibliotheken und Rechenzentren.

Methode der Wahl waren i.A. online-Fragebögen. Die Top-Antworten aus diesen Fragebögen zu gewünschten Serviceleistungen hat Kronenwett (2016) in ihrer Masterarbeit in einer Tabelle zusammengefasst (s. Tab.1, abgewandelt und ergänzt). In zwei Fällen wurden nachträglich zur Ergänzung der online-Umfragen und zur Klärung von bei der Auswertung aufgetauchten Fragen Interviews durchgeführt (Humboldt-Universität zu Berlin, Simukovic et al. 2014; Leibniz Universität Hannover, Hauck, 2016). Diese bestätigen und vertiefen die Ergebnisse aus den online-Umfragen. Weitere ergänzende Interviews sind geplant (z.B. Universität Trier: Lemaire et al., 2016; Christian-Albrechts-Universität zu Kiel: Paul-Stüve, 2015; Technische Universität Darmstadt: Langhanke & Stille, 2015). An den Hochschulen in Baden-Württemberg (Tristram, 2015: Mannheim, Heidelberg, Karlsruhe, Stuttgart, Hohenheim, Tübingen, Ulm, Freiburg, Konstanz) wurden nur Interviews geführt, ohne vorherige online-Befragung.

In allen Befragungen zeigen sich große Unterschiede in Bewusstsein und Vorbildung zum FDM sowie bei dessen Umsetzung. Demzufolge waren gewünschte Serviceleistungen vor allem Beratung, einschließlich Schulungen (s. Tab.1) sowie Best Practice Beispiele. In der Graduiertenausbildung ist das Thema i.A. angekommen. Speicherplatz für die operationelle Speicherung der Daten, die den Austausch mit Projektpartnern erlaubt (Datenbox), sowie für die Langzeitspeicherung waren ebenfalls

Thema (s. Tab.1). Die Angebote sollten fach- und datenspezifisch sein. Die angebotenen Lösungen sollten einen geringen Aufwand erfordern und der eigene Nutzen des FDM sollte erkennbar sein. Außerdem müssen Verantwortlichkeiten für die Daten an den einzelnen Institutionen definiert werden. Teilweise wurde gewünscht, dass bereits vorhandene Services besser publik gemacht und verständlicher erklärt werden.

Tabelle 1: Befragungsergebnisse zu gewünschten Serviceleistungen an anderen deutschen Hochschulen. Tabelle übernommen, abgewandelt und ergänzt aus Kronenwett (2016, Tab.11 S. 88).

Hochschule	Publikation	Gewünschte Serviceleistungen, Top-Antworten
Humboldt-Universität zu Berlin	Simukovic et al. 2013	Speicherplatz, rechtliche und technische Beratung
Westfälische Wilhelms-Universität Münster	Herwig et al. 2014	Rechtliche, technische und allgemeine Beratung
Philipps-Universität Marburg	Krähwinkel, 2015	Technische, rechtliche und allgemeine Beratung
Christian-Albrechts-Universität zu Kiel	Paul-Stüve, 2015	Speicherplatz, technische und rechtliche Beratung
Technische Universität Darmstadt	Langhanke & Stille, 2015	„Gesichertes und verlässliches Datenarchiv“, rechtliche und technische Beratung, außerdem: TU- oder Hessen-Datenbox
Leibniz Universität Hannover	Hauck et al., 2016	Rechtliche und technische Beratung, Speicherplatz
Universität zu Köln (Philosophische Fakultät)	Kronenwett, 2016	Rechtliche und technische Beratung, Speicherplatz
Universität Trier	Lemaire et al., 2016	Infrastrukturen zur Archivierung, allgemeine, technische und rechtliche Beratung
Technische Universität Hamburg	Feldsien-Sudhaus & Rajski, 2016	Rechtliche und technische Beratung, sichere TU- oder Hamburg-Datenbox, gesichertes und verlässliches Datenarchiv

Jede Hochschule hat eigene fachliche Schwerpunkte und eigne Voraussetzungen in Bezug auf FDM. Die Befragung an der FSU dient dazu, Serviceangebote auf konkrete Bedürfnisse für Projekte an der FSU zuschneiden zu können. Dabei soll berücksichtigt werden, welche Anteile des FDM so projektspezifisch sind, dass sie nur innerhalb dieser Projekte in dedizierten Teilprojekten geleistet werden können und welche Serviceanteile effizienter durch eine zentrale Struktur erbracht werden können. Auch fachspezifische und universelle Angebote sollen unterschieden werden. Außerdem dient die Befragung der Sichtbarmachung der bereits angebotenen und geplanten Services sowie der Kontaktaufnahme und der Eröffnung der Kommunikation mit den Verantwortlichen der Projekte.

1.4 Ziel

Ziel ist eine detaillierte Bedarfserhebung bei laufenden Verbundprojekten. Aus strategischer Sicht ist eine institutionalisierte Unterstützung des FDM in Verbundprojekten in gleich mehrerlei Hinsicht vorteilhaft: Durch die zentralisierte Bereitstellung von Beratungsangeboten aber auch FDM-

Plattformen für Verbundprojekte werden bessere Skalierbarkeit und Einsparungseffekte (Betrieb weniger statt vieler Plattformen, gebündelte Expertise) erreicht. Gleichzeitig findet im Rahmen von Verbundprojekten eine enge Zusammenarbeit mit den außeruniversitären Forschungseinrichtungen am Standort Jena statt. Hier können weitere Synergien zur Steigerung von Effektivität und Effizienz des FDM genutzt werden. Im Ergebnis wird eine gesteigerte Qualität der Datenhaltung und daraus folgend eine höhere Nachnutzbarkeit der Daten erwartet.

Hierzu sind geeignete Prozessmodelle zu entwickeln, die einen nahtlosen Übergang zwischen unterschiedlichen Phasen des Lebenszyklus gewährleisten. Zudem hat die durch die AG FDM an der FSU durchgeführte Statusanalyse Bedarf zum Aufbau zusätzlicher Angebote in unterschiedlichen Phasen des Lebenszyklus ergeben: Dies sind eine Professionalisierung der Beratung in der Antragsphase sowie Unterstützung bei Entscheidungen rund um die Bewahrung von Forschungsdaten nach Projektende und der Aufbau einer institutionalisierten technischen Unterstützung des Datenmanagements während der Projektlaufzeit.

Mittels einer Analyse geplanter, laufender und abgeschlossener Verbundprojekte sowie einer strukturierten Befragung der Verantwortlichen in diesen Vorhaben soll erhoben werden, welche Aufgaben im Bereich des FDM im Detail anfallen und wie diese dort gelöst werden. Insbesondere soll ermittelt werden, welche Anteile des FDM so projektspezifisch sind, dass sie nur in dedizierten Teilprojekten geleistet werden können, und welche Serviceanteile existieren, die effizienter durch eine zentrale Struktur erbracht werden könnten. Die Befragung soll Grundlage für die Planung von weiteren Schritten zum institutionalisierten FDM an der FSU sein. Mit Blick auf die Entwicklung des zentralen Service sollen dabei die Bereiche Infrastruktur, Beratung und Schulung berücksichtigt werden.

Aufgrund der besonderen Bedeutung der Biodiversitätsforschung für die FSU (s. Profillinien) und die ausgewiesene Expertise im FDM für diese Disziplin, sollen Verbundprojekte aus diesem Bereich als Piloten für eine Umsetzung dienen. Die entwickelten Konzepte sollen aber auf andere Disziplinen übertragbar sein.

2 Methoden

Die Befragung wurde anhand von Interviews umgesetzt. Im Gegensatz zum online-Fragebogen ist es hier einfacher konkrete Projekte zur Teilnahme zu bewegen. Wichtige Vorteile sind auch der persönliche Kontakt zu den Befragten sowie die Möglichkeit der Befragten, Antworten zu erklären und eigene Punkte zu ergänzen.

2.1 Projektauswahl

Um eine Auswahl an Verbundprojekten der FSU treffen zu können wurde zunächst eine Liste aller an der FSU laufender Forschungsprojekte in koordinierten Programmen der DFG (Forschungszentren, Sonderforschungsbereiche, Forschergruppen, Schwerpunktprogramme) und des BMBF (Kollegs, Partnerschaften) erstellt. Außerdem wurden Projekte in die Auswahlliste aufgenommen, die nicht länger als zwei Jahre beendet waren. Zusätzlich gibt es an der FSU EU-geförderte Verbundprojekte. Es war allerdings nicht möglich, sich im Vorfeld der Befragung einen Überblick über diese Projekte mit den entsprechenden Eckdaten (Förderzeitraum, beteiligte Fachbereiche, Sprecher etc.) zu verschaffen. Deshalb wurden diese für die Auswahlliste nicht berücksichtigt.

Aus der Liste mit 35 Projekten wurden dann die Projekte so ausgewählt, dass ein möglichst breites Spektrum an Laufzeitabschnitten (in Planung, Beginn, Mitte, Ende, beendet, s. Abb. 1) und Disziplinen abgedeckt wurde. Zudem wurde darauf geachtet, Projekte zu befragen, von denen bereits bekannt war, dass sie ein INF-Projekt oder Analogon dazu enthalten sowie Projekte, die nicht über ein solches

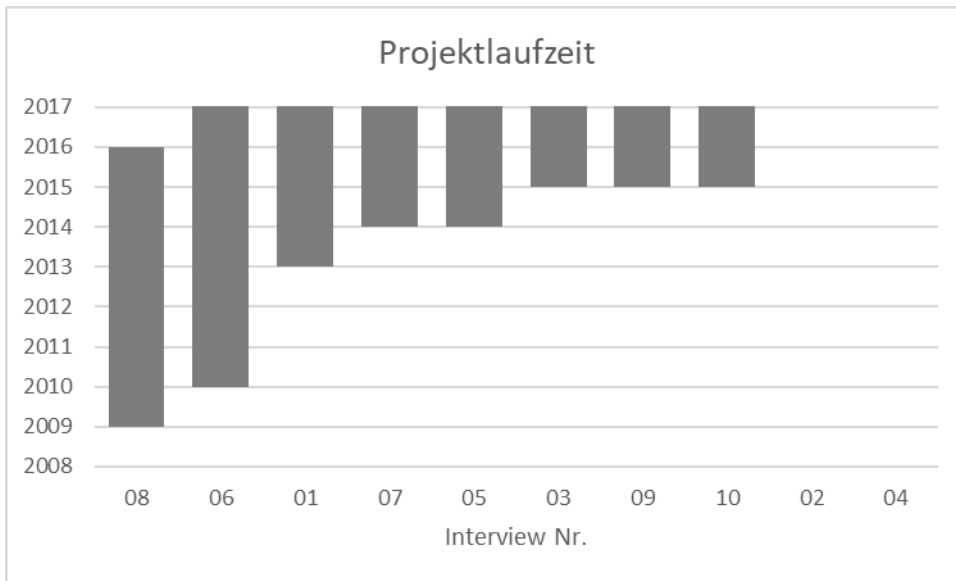
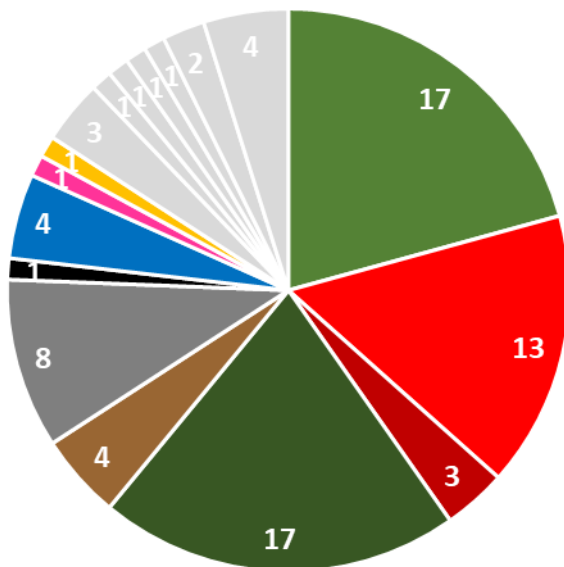


Abbildung 1: Stand der Projektlaufzeit der ausgewählten Projekte. Das Projekt mit der Interviewnummern 02 hat die Projektskizze eingereicht, das mit der Nummer 04 hat die Fördermittelzusage erhalten.

Anteil Disziplinen Verbundprojekte FSU



Anteil Disziplinen ausgewählte Projekte



Abbildung 2: Anteil der Disziplinen in allen laufenden und kürzlich beendeten Verbundprojekten der FSU (links) im Vergleich zum Anteil der Disziplinen in den ausgewählten Projekten (rechts) (Stand Mai 2017). Die Zahlen geben die Anzahl der Beteiligungen der jeweiligen Disziplin an. An einem Projekt können mehrere Disziplinen beteiligt sein (Mehrfachnennungen möglich). *Biologie, Medizin, Pharmazie, Chemie, Geowissenschaften, Physik, Astrophysik, Materialwissenschaft, Psychologie, Geschichte, nicht befragte Bereiche (hellgrau, im Uhrzeigersinn: Soziologie, Orientalistik, Kunstgeschichte, Literaturwissenschaft, Philosophie, Mathematik, Informatik).*

Teilprojekt verfügen. Auf die Befragung von Projekten an denen die Arbeitsgruppe König-Ries selbst mit INF-Teilprojekten beteiligt ist, wurde bis auf eine Ausnahme (Befragung diente dem Test des Fragebogens) ebenfalls verzichtet da hier wenig neue Erkenntnisse zu erwarten waren. Im Ergebnis wurden sieben Projekte mit DFG-Förderung befragt: vier Sonderforschungsbereiche (SFB) und drei Forschergruppen sowie drei BMBF-geförderte Projekte: ein Forschungscampus, ein Kolleg und eine Partnerschaft für Innovation.

Der insgesamt relativ hohe Anteil an Biowissenschaften (Biodiversität, Mikrobiologie), Chemie und Medizin an den Verbundprojekten insgesamt spiegelt sich auch in der fachlichen Zusammensetzung der zur Befragung ausgewählten Projekte wieder (Abb. 2). Nach Anzahl an Professuren beträgt der Anteil an Naturwissenschaften an der FSU gut ein Drittel (38%), der vom Fachbereich Medizin etwa 21% (Stand Dezember 2016). Obwohl der Anteil an Geisteswissenschaften nach Studierendenzahlen knapp ein Drittel ausmacht (30%), waren es bei den Verbundprojekten nur 10%. Hier wurde ein Projekt zur Befragung ausgewählt. Rechtswissenschaften und Wirtschaftswissenschaften, die nach Anzahl der Professuren etwa 10% der FSU darstellen, waren bei den Verbundprojekten nicht vertreten.

2.2 Interviewleitfaden

Der Interviewleitfaden (s. Anhang) basiert auf den Fragebögen von Simukovic et al. 2013 und Bauer et al. 2015 sowie dem Interviewleitfaden von Simukovic et al. 2014. Aus den Fragebögen wurden die Kategorien „Datentypen, -formate und -größe“, „Datenarchivierung und -sicherung“, „Zugänglichkeit und Nachnutzung“ sowie „Infrastruktur und Service“ übernommen und angepasst. Eine Kategorie „Lehre“ wurde ergänzt. Insgesamt enthält der Leitfaden 46 Fragen. Zum Einstieg diente die Frage „Haben Sie, bevor wir uns bei Ihnen gemeldet haben, schon etwas von der Kontaktstelle Forschungsdatenmanagement der Friedrich-Schiller-Universität gehört“. Außerdem wird der Befragte zunächst gebeten, das Projekt vorzustellen, um in den Redefluss zu kommen. Der Leitfaden wurde dem Interviewpartner im Vorfeld zur Verfügung gestellt.

2.3 Durchführung der Befragung

Anhand des Leitfadens wurden Verantwortliche der zehn ausgewählten Projekte in teilstrukturierten Interviews befragt. Diese dauerten zwischen 45 und 90 Minuten. Es wurde auf Themen der Befragten eingegangen, die nicht im Leitfaden vorgesehen waren und nur zur Ergänzung auf den Leitfaden zurückgegriffen. Diese Vorgehensweise stellt einen Kompromiss dar, der möglichst umfassende und gleichzeitig möglichst vergleichbare Antworten erlauben soll. Um die Inhalte des Leitfadens im Blick zu behalten und gleichzeitig flexibel auf die Befragten reagieren zu können, wurden die Interviews in der Regel von zwei Personen durchgeführt.

Befragt wurden Projektverantwortliche wie Sprecher (fünf Projekte, davon einer sehr datenaffin), Teilprojektleiter (drei Projekte) und Datenmanager (zwei Projekte). Die meisten Interviews fanden am Arbeitsplatz des Interviewpartners statt. Zur späteren Analyse wurden die Interviews per Audioaufnahme aufgezeichnet und dann von wissenschaftlichen Hilfskräften mit f4 transkribiert (Version 3.1.0, 2007, dr. dresing & pehl GmbH, Marburg). Die Antworten aus den Befragungen wurden zunächst pro Interview und in einem zweiten Schritt als Vergleich zwischen den Interviews zusammengefasst.

3 Ergebnisse

3.1 Aktueller Stand zum Forschungsdatenmanagement

In sieben Projekten gaben die Befragten an, sich bereits bei der **Projektplanung** mit FDM auseinandergesetzt zu haben. In einem Projekt aus den Geisteswissenschaften wurde angegeben, dass es im Projekt kaum Bedarf an einem gemeinsamen FDM gibt, da die Datenmengen im Megabyte-Bereich liegen und die Wissenschaftler sehr individuell arbeiten. In einem Projekt aus dem Bereich Physik wurde sich im Vorfeld nicht explizit mit FDM befasst, in den Teilprojekten werden jedoch einzelne Aspekte des FDM (z.B. gesicherte und redundante Datenhaltung) organisiert und umgesetzt. Beim ältesten der befragten Projekte, welches inzwischen beendet ist, gab es während der Beantragung keine Planung zum FDM, allerdings ist ein strukturiertes FDM für ein Folgeprojekt vorgesehen.

Die **Umsetzung des FDM** war sehr heterogen. In sieben Projekten waren Data Policies auf Projekt- oder Institutionsebene vorhanden. Häufig wurde sich auf die Empfehlungen zur Sicherung guter wissenschaftlicher Praxis berufen (DFG 1998, 2013). In zwei Projekten war der Umgang mit Forschungsdaten und die entsprechenden Verantwortlichkeiten sehr stark geregelt und wurde durch sogenannte „Briefings“ sowie regelmäßige Kontrollen durchgesetzt. In beiden Projekten wird sehr industrienah bzw. mit Industriepartnern geforscht. Andere Gesprächspartner gaben an, dass die praktische Umsetzung durch die ausführenden Wissenschaftler mitunter nicht den Vorgaben entspricht bzw. zu Gunsten anderer, als wichtiger erachteter Tätigkeiten auf ein Minimum beschränkt wird. Insbesondere bezüglich der Langzeitspeicherung und Nachnutzung der Daten sind die Policies noch zu wenig konkret.

Die 2015 gegründete **KS FDM** war sechs Befragten vor der Kontaktaufnahme für die Interviews noch nicht bekannt. Bei den anderen Projekten gab es entweder eine direkte Beteiligung der AG König-Ries oder es bestanden andere Kontakte z.B. über das Michael Stifel Center Jena.

3.2 Charakterisierung der Daten

Bei den genannten **Quellen** der generierten Forschungsdaten zeigte sich das zu erwartende breite Spektrum. Es handelte sich vor allem um Beobachtungen, Experimente und Messungen (z.T. Sensordaten), bildgebende Verfahren (Mikroskopie, Spektroskopie, Spektrometrie, Kristallographie), Sequenzen (DNA, RNA, Proteine) und häufig auch Simulationen (fünf Projekte). Es wurden aber auch Textmining, Archivmaterialien (Textdokumente, Bilder, Ton- und Videoaufzeichnungen) und Umfragen genannt. Grundlage für die Datenerhebung waren außerdem biologische und geologische Proben sowie chemische Substanzen.

Das geschätzte **Datenaufkommen** der befragten Projekte reicht von wenigen Megabyte (Word- und Excel-Dokumente) bis zu einigen Petabyte (Lichtblatmikroskopie). In den meisten Projekten bilden digitale Daten den weitaus größten Teil der anfallenden Forschungsdaten. Handschriftliche Notizen und analoge Laborbücher waren zum Teil noch üblich. Vier der Befragten gaben aber an, dass eine Umstellung auf digitale Laborbücher angestrebt wird oder bereits in Planung ist.

Auch die anfallenden **Datenformate** zeigten ein breites Spektrum. Textdokumente und tabellarische Daten in den Formaten der gängigen Office-Produkte wurden in allen Projekten verwendet. In neun Projekten fielen auch Bilddateien an und in Projekten mit Simulationen auch Modelle und Quellcode. Zweimal wurden auch Videos genannt, einmal Audiodateien und zweimal relationale Datenbanken. Häufig wurden allgemeine oder fachspezifische Standardformate verwendet. Die Daten, die mit Messgeräten erhoben wurden oder von bildgebenden Verfahren stammen, lagen aber meist geräte- oder sogar herstellerspezifisch vor. Die Überführung dieser Formate in Standardformate für die Langzeitspeicherung kam nur in Ausnahmefällen vor. In zwei Fällen wurden auch eigene Formate geschaffen. Sensible bzw. vertrauliche Daten kamen in sieben Fällen vor, dazu gehörten Patienten- bzw. personenbezogene Daten (vier Projekte), patentrechtlich relevante Daten (drei Projekte) sowie andere sensible/vertrauliche Daten (ein Projekt).

In einigen Projekten findet eine inhaltliche **Qualitätssicherung** der Daten (Plausibilitätsprüfung) schon vor oder während der Analyse statt, zum Teil an den Messgeräten selbst, zum Teil durch den erhebenden Wissenschaftler oder andere Spezialisten (vier Projekte). Das ist in einigen Bereichen schon etablierter Standard (z.B. next generation sequencing), findet in Form von Stichproben statt (ein Projekt) oder die Umsetzung und z.T. auch Automatisierung wird angestrebt und entwickelt (mindestens zwei Projekte). Automatisiert ist v.a. die technische Qualitätssicherung (z.B. Überprüfung auf fehlende Daten, Syntaxfehler). Diese wird meistens dann umgesetzt, wenn Daten vom Messgerät auf größere Speicherformen (z.B. Netzlaufwerke in den Arbeitsgruppen) übertragen werden.

3.3 IT-Infrastruktur und Service

Wo es die generierte Datenmenge erlaubt und keine anderweitigen Regelungen bestehen, **speichern** Wissenschaftler Daten in der Projektphase bevorzugt auf dem persönlichen Dienstrechner (z.T. aber auch auf privaten Rechnern, v.a. Studenten). Bei Messungen an Großgeräten und bildgebenden Geräten werden die Daten im Allgemeinen auf den Rechnern an den Geräten belassen (mindestens zwei Projekte). Neun der Projekte nutzen eine oder mehrere Formen von Netzlaufwerksystemen (z.B. NAS, SAN) für die operationelle Speicherung der Daten bzw. sind dabei, für Daten von bestimmten Messgeräten dezidierte Plattformen zum Datenmanagement einzurichten. Bei den Speichersystemen handelt sich mehrheitlich um Lösungen, die am jeweiligen Institut oder in einer Arbeitsgruppe vorhanden sind. Die Daten werden häufig in mehreren Kopien und z.T. auch in mehreren Versionen vorgehalten.

Verantwortlich für die Speicherung der Daten sind überwiegend die Wissenschaftler selbst. Zum Teil wird aber auch der (Teil-)Projektleiter in der Verantwortung gesehen, der in einigen Fällen auch Überprüfungen durchführt. Teilweise gibt es IT-Personal, das aber in erster Linie für klassische IT-Aufgaben verantwortlich ist und FDM nur unterstützend begleitet. Dezidierte Datenmanager gibt es in Projekten mit einem INF-Projekt. In einem Institut soll eine W3-Professur Bioinformatik geschaffen werden, die auch Datenmanagement Expertise für verschiedene Projekte im Haus anbieten soll.

Von den Servern und z.T. auch den dienstlichen Rechnern werden regelmäßig **Backups** erstellt. Zwei Projekte gaben an, den Backup-Service des URZ zu nutzen. Nur im Projekt aus den Geisteswissenschaften (Daten im Megabyte-Bereich) wurden die Daten manuell auf USB-Speichermedien gesichert. Insbesondere die Projekte, an denen medizinische Einrichtungen (z.B. UKJ) und außeruniversitäre Forschungsinstitute (z.B. MPI, Fraunhofer Institut, FLI) beteiligt sind, steht eine sehr gute, dem Bedarf entsprechende technische Infrastruktur zur Verfügung.

Die **Dienste des URZ** werden in den befragten Projekten sehr selektiv genutzt und sind zum Teil auch nicht bekannt. Dies ist vor allem darauf zurückzuführen, dass insbesondere bei Projekten, an denen das UKJ und außeruniversitäre Forschungseinrichtungen beteiligt sind, eigene IT-Abteilungen und entsprechende Dienste bestehen und genutzt werden. Zum Teil wurde sich bewusst gegen eine zentrale (universitätsverwaltete) Infrastruktur entschieden (drei Projekte). Gründe dafür sind vor allem Sicherheitsbedenken (z.B. im Zusammenhang mit patentrechtlichen Bestimmungen oder Datenschutzbestimmungen) bzw. das Bedürfnis die Datenhoheit zu behalten. Zum anderen wurden lange verwaltungstechnische Wege genannt. Für ein Projekt werden Server im URZ gehostet. Ein weiteres nutzt den File-Upload-Service zum Versenden größerer Dateien. Der Backup-Service (Bandserver) wird in zwei Projekten genutzt.

Als **Tools**, die zur Unterstützung der Projekte in der Kommunikation und im Datenmanagement genutzt werden, wurden genannt: Projektmanagement (drei Projekte), Kommunikation / Kollaboration (ein Projekt: interne Cloud-Dienste, SharePoint, Tempo Box), elektronische Laborbücher / Labormanagement (zwei Projekte: LIMS, ein Projekt: E-Notebook, zwei Projekte: andere), Datenmanagement (zwei Projekte: BEXIS 2, je ein Projekt: SEEK, ChemBioSysDB, Canto Cumulus), rechtssichere Dokumentation (ein Projekt), Versionierung (ein Projekt), Literaturverwaltung (nur von einem Projekt genannt, obwohl möglicherweise in allen genutzt). Als sinnvoll angesehen werden Tools zur Vorhaltung und Sicherung von Daten mit Zugriffskontrolle sowie Strukturen zur Analyse von Daten auf "höherer Ebene" mit teils noch unbekanntem Fragestellungen (ein Projekt) und Versionierungstools (ein Projekt). Kein unmittelbarer Bedarf an solchen bzw. weiteren Tools wurde in drei Projekten geäußert. Von einem Befragten wurde angesprochen, dass für einzelne Probleme bereits **Softwarelösungen** existieren, diese aber erst zu einem funktionierenden Gesamtsystem zusammengeführt werden müssen.

3.4 Langzeitarchivierung, Nachnutzbarkeit und Publikation von Daten

Bei den meisten Projekten verbleiben die Forschungsdaten **nach Projektende** an den jeweiligen Instituten, auf den Servern, auf Festplatten oder auf CDs/DVDs (letztere beiden Medien v.a. bei Abschlussarbeiten). In zwei Projekten ist noch nicht geklärt, was mit den Daten nach Projektende passiert. Davon ist eins aber noch in der Vorbereitungsphase und misst dem Nachhaltigkeitskonzept einen hohen Stellenwert bei. In einem Projekt aus den Geisteswissenschaften werden nur die wissenschaftlichen Publikationen (Aufsätze) als relevant für andere Wissenschaftler angesehen, nicht die Daten an sich. Deshalb behält jeder Wissenschaftler die von ihm generierten Daten (Exzerpte, Notizen, Tabellen etc.) nur auf seinem eigenen Rechner und seinem persönlichen USB-Laufwerk.

Hinsichtlich der **Langzeitspeicherung** der Daten ist allen Befragten die „10-Jahre-Regel“ aus den DFG Empfehlungen zur Sicherung guter wissenschaftlicher Praxis (1998, 2013) bekannt. Dagegen war die Ende 2015 von der DFG herausgegebenen Leitlinien zum Umgang mit Forschungsdaten kaum bekannt, genauso wie die Ende 2016 verabschiedete Leitlinie der FSU zum FDM. Die Umsetzung der „10-Jahre-Regel“ erfolgt bei den meisten Projekten pragmatisch und beschränkt sich auf das einmalige Sichern zum Projektende, zum Ende einer Qualifizierungsarbeit oder im Zusammenhang mit einer wissenschaftlichen Publikation. Fast alle Befragten betrachten die eigene Praxis der Langzeitspeicherung jedoch kritisch und sehen Bedarf für Verbesserung. Als Ursachen für diese Umstände gaben die Befragten die damit verbundenen Kosten für Speicherplatz an sowie fehlendes Personal für Datenmanagement und langfristige Kuratierung, aber auch mangelnde Bereitschaft einzelner Wissenschaftler, die Daten für die Langzeitspeicherung aufzubereiten und zu archivieren oder auch der geringe Wert der Daten für andere. Bei einem Projekt werden Rohdaten an Geräten für zehn Jahre gespeichert und danach gelöscht. Als Grund dafür wird der begrenzte Speicherplatz angegeben sowie geräte- und herstellerspezifische Formate, die erwarten lassen, dass diese nach zehn Jahren nicht mehr lesbar sind. Bei einigen Projekten und in bestimmten Anwendungsfällen (bestimmte Daten: z.B. Gensequenzen, neue Substanzen) werden die Daten insbesondere im Zusammenhang mit wissenschaftlichen Publikationen in fachspezifischen Datenbanken oder Repositorien abgelegt (mindestens drei Projekte). Dies geschieht aufgrund von Vorgaben der Fach-Community bzw. der Fördermittelgeber. In zwei Projekten werden auch die physischen Proben nach Community-Standards archiviert (einmal lokal im Institut und einmal in einem nationalen Probenarchiv).

Die **Erfassung von Metadaten** ist ebenfalls sehr heterogen, sowohl zwischen den befragten Projekten als auch innerhalb der Projekte (je nach Teilprojekt, Wissenschaftler, Art der Daten). In drei Projekten wird die standardisierte Dokumentation der Erstellung und Verarbeitung der Daten als essentiell hervorgehoben, in drei anderen wird sie als rudimentär und individuell unterschiedlich beschrieben. Verantwortlich sind im Allgemeinen die Wissenschaftler, die die Daten erheben. Wird an Großgeräten gemessen bzw. werden bildgebende Geräte genutzt (mindestens vier Projekte), erfolgt in der Regel eine automatische Erfassung einiger wichtiger Metadaten (z.B. Gerät, Datum, Uhrzeit, z.T. Nutzer). Werden Labormanagementsysteme oder elektronische Laborbücher genutzt (sechs Projekte), wird die Entstehung und Verarbeitung der Daten größtenteils dort einheitlich beschrieben. Zum Teil liegen eigene Metadatenschemata vor (zwei Projekte). Von einem Befragten wird die Bedeutung der Auffindbarkeit von Daten innerhalb des Projektes durch gute Metadaten hervorgehoben. Beim Hochladen in Datenbanken / Repositorien werden die Metadaten nach den dort vorgegebenen Standards ergänzt.

Der **Datenaustausch** erfolgt zwischen den Projektpartnern entweder über die eingesetzten Datenmanagementtools (z.B. zwei Projekte: BEXIS 2, ein Projekt: SEEK, ChemBioSysDB), über Zugriffsrechte direkt auf den Projektservern (sieben Projekte), über FTP Dienste (zwei Projekte) oder per E-Mail (vier Projekte). Zum Teil werden auch Speichermedien (Festplatten, CDs/DVDs) verschickt (mindestens ein Projekt). Nach außen werden Daten in allen Projekten nur auf direkte Anfrage

gegeben. Ausnahmen hierbei stellen einerseits Community-Datenbanken dar (Datenupload im Rahmen einer wissenschaftlichen Publikation, s.o. Langzeitspeicherung), andererseits sensible / vertrauliche Daten (keine Herausgabe). Ein Interviewpartner gab an, dass bei einer Anfrage zu Daten, zu denen noch keine wissenschaftliche Publikation erschienen ist, die Möglichkeit einer gemeinsamen wissenschaftlichen Veröffentlichung eingeräumt wird. Bei dem befragten geisteswissenschaftlichen Projekt findet üblicherweise kein Datenaustausch statt (s.o. Projektende). Bei dem Projekt, das sich noch in der Vorbereitungsphase befindet, müssen die Austauschmodalitäten noch im Rahmen des INF-Projektes geklärt werden.

Von der Möglichkeit Daten als eigenständige **Publikation** unabhängig von einer wissenschaftlichen Veröffentlichung zu publizieren wird in den befragten Projekten nur vereinzelt Gebrauch gemacht. In Fachbereichen, in denen etablierte internationale Datenbanken bestehen (z.B. NCBI GEO, GeneExpress am EBI), werden diese sowohl für die Publikation der eigenen Daten, als auch für die Nachnutzung fremder Daten genutzt. Einige der Befragten sehen es als wichtige Aufgabe an, Forschungsdaten barrierefrei zugänglich zu machen, inklusive guter Metadaten und Qualitätssicherung. Wenn noch keine eigene wissenschaftliche Publikation erschienen ist, gibt es Sperrfristen von zwei bis fünf Jahren auf Daten.

Hinsichtlich der zunehmend geforderten **Nachnutzbarkeit** von Forschungsdaten gaben einige der Befragten an, dass dies durch die wissenschaftlichen Publikationen bzw. durch die genutzten Fachdatenbanken gewährleistet sei. Oft wird angegeben, dass die Daten hierdurch nach außen ausreichend sichtbar sind (fünf Projekte). Konkrete Pläne zur Förderung der Nachnutzbarkeit gibt es nicht, werden aber in zwei Projekten eingehend diskutiert. Nur ein Befragter gab an, dass eine gegenseitige Nachnutzung von Daten in der Community üblich ist. In einem weiteren Projekt ist eine eigene Datenbank zum Teilen bestimmter Daten angedacht, die als Grundlage für weitere Publikationen der Fachgemeinschaft dienen können. Hier gibt es bereits ähnliche Datenbanken bei einzelnen anderen Gruppen der Community, aber kein gemeinsames Repositorium. Bei einem der Projekte wurde konkret angegeben, dass normalerweise keine Nachnutzung stattfindet.

3.5 Lehr- und Kursangebote

Die **Vorbildung** neuer Mitarbeiter bezüglich FDM wird als sehr unterschiedlich beschrieben, z.T. in Abhängigkeit von der fachlichen und institutionellen Herkunft. Besteht z.B. ein Hintergrund im Fachbereich Informatik, sind die Kenntnisse sehr gut. Ansonsten sind sie wenig bis mittelmäßig vorhanden. Die Sensibilisierung für das Thema nimmt aber zu. In drei der befragten Projekte werden Softskills im Bereich FDM gewünscht oder sind Voraussetzung.

Zwei der Befragten gaben an, dass FDM-Themen bereits in bestehenden **Veranstaltungen** der grundständigen Lehre enthalten sind (z.B. Empiriepraktikum, Vorlesungen zum wissenschaftlichen Arbeiten etc. für Bachelor- und Masterstudenten). FDM-Kurse an Graduiertenschulen werden von zwei Befragten als wirksam angesehen. Workshops zum Datenmanagement für gestandene Wissenschaftler werden in zwei Projekten angeboten. In fünf Projekten werden die Mitarbeiter in Briefings und über ihre direkten Vorgesetzten über die Modalitäten beim Umgang mit Daten informiert. Die Ansiedlung der Inhalte im Bachelorstudium wird von zwei Befragten als zu früh angesehen, da die Studenten bis dahin noch nicht mit Daten gearbeitet haben. Begleitend zur Bachelorarbeit wird es aber im Rahmen der Betreuung als sinnvoll erachtet. Die Anleitungen zum FDM sollten zeitlich nicht zu aufwendig sein und könnten ggf. auch im Rahmen von Praktika vermittelt werden (Methoden, digitale Laborbücher etc.).

Generell wird ein **Bedarf** gesehen, Studenten und Mitarbeitern Kurse zum FDM anzubieten (vier Befragte). Ein zentrales Angebot der FSU wird dabei begrüßt (drei Befragte). Dabei sollte aber darauf geachtet werden, dass der Mehrwert von FDM von den Teilnehmern erkannt wird und sich Mitarbeiter

nicht bevormundet fühlen. Demgegenüber wurden aber auch Bedenken geäußert, dass FDM fachspezifisch sehr unterschiedlich ist, was eine übergreifende Veranstaltung schwer vorstellbar macht (zwei Befragte). Um auf konkrete Probleme des Fachs eingehen zu können, müsste jeder der Lehrenden in seinem Fach auf FDM eingehen. Nur der Interviewpartner aus den Geisteswissenschaften äußerte, dass in seinem Fachgebiet kein Bedarf an Veranstaltungen zum FDM besteht.

3.6 Anmerkungen und Wünsche

Wenn generelle Probleme angesprochen und Wünsche zur Unterstützung geäußert wurden, waren diese sehr ähnlich. Immer wieder tauchten dabei die Begriffe Finanzen, Personal, bessere Workflows und Schnittstellen, Automatisierung und Verstetigung auf. Von einem Befragten wurde explizit begrüßt, dass die Hochschulleitung der FSU FDM als zentrale Aufgabe betrachtet und konkrete Maßnahmen ergriffen hat.

Finanzielle Mittel wurden unter anderem für Software-Lizenzen (z.B. elektronische Laborbücher) gewünscht. Ein Hinwirken auf einen größeren Lizenznehmerkreis an der FSU und eine Abdeckung über Landeslizenzen wurde dabei als sinnvoll erachtet. Von einigen Befragten wurden mehr finanzielle Mittel für spezielle Hardware (z.B. Feldrechner) angesprochen. Generell schien die Ausstattung und Beschaffung von Hardware bei den befragten Projekten kein Problem zu sein.

Personal zur administrativen Betreuung von Servern und Software sowie zur Migration von Datenbanken mit unterschiedlichen Strukturen oder von Daten auf CDs/DVDs in modernere Speichermedien waren häufig genannte Punkte. Dafür könnte es einen zentralen Service geben (zwei Befragte) oder Personal in den einzelnen Instituten zur Verfügung stehen (vier Befragte). Momentan ist IT-Personal, wenn überhaupt vorhanden, mit den klassischen administrativen Aufgaben ausgelastet. Dazu ist ggf. eine Umstrukturierung / Zusammenarbeit von Instituten nötig. Ein Vorschlag war, dass Doktoranden der Informatik über den Aufbau, den Betrieb und die Anpassung fachspezifischer Anwendungen promovieren.

Ein weiterer mehrfach geäußelter Wunsch war eine Möglichkeit zur Sicherung und Organisation von Dateien auf Projektebene, z.B. durch eine **FSU-Cloud für Projekte**. Dabei sollte der Zugriff für Universitätsangehörige auch aus dem Ausland sowie für außeruniversitäre und ausländische Projektpartner möglich sein (drei Befragte). Ziel ist u.a., dass der einzelne Wissenschaftler nicht mehr seine Daten auf dem eigenen Rechner vorhält, sondern alle Projektteilnehmer Zugriff haben und vor allem Arbeitsschritte eingespart werden (Speichern auf eigenem Rechner, Übertragung auf Server für Projektlaufzeit, Übertragung in Langzeitarchiv). Die Cloud soll von einem vertrauenswürdigen Host (URZ, Institut) verantwortet, kontrolliert und gesichert werden. Für einen Befragten wäre auch eine Cloud auf nationaler oder EU-Ebene denkbar (Stichwort NFDI oder EOSC).

Ein weiteres Thema war **Speicherplatz**, wobei die Meinungen auseinandergingen, ob dieser zentral zur Verfügung gestellt werden sollte (zwei Befragte) oder ob die einzelnen Institute dieses Thema selbst lösen sollten (ein Befragter). In einem Projekt wird auch der Bedarf gesehen, das bereits vorhandene Git-Repository vom URZ hosten und administrieren zu lassen.

Ein wichtiger Punkt war **Automatisierung**, um den Aufwand für den einzelnen Wissenschaftler zu minimieren (Speicherung zur Bearbeitung auf eigenem PC, Übertragung auf Server zur Sicherung und gemeinsamen Nutzung während der Projektlaufzeit durch die Projektpartner, Speicherung im Repository für Langzeitarchivierung und Publikation, ggf. Attributierung). So wurde z.B. ein automatischer Upload von plausiblen Rohdaten in ein Backup-System (ein Befragter) bzw. in ein Langzeitarchiv gewünscht (ein Befragter). Bei Letzterem sollten natürlich die Sperrzeiten für die Freigabe gegenüber (externen) potentiellen Nutzern berücksichtigt werden. Ebenso wurde ein digital-

protokollierter (standardisierter) Prozess zur Dokumentation der Experimente an den Geräten in einem digitalen Logbuch angesprochen (ein Befragter).

Als notwendig erachtet, insbesondere wenn die angebotene technische Infrastruktur genutzt werden soll, wurden auch **Schulungen** (vier Befragte) **sowie Beratung** (drei Befragte), z.B. für die zu nutzenden Tools, aber auch zum sensiblen Umgang mit Daten / zur Datensicherheit, zu technischen und rechtlichen Fragen, insbesondere in Bezug auf die Projektplanung und -konzeption und die Antragstellung bei einem potentiellen Fördermittelgeber sowie zu Langzeitarchivierungsstrategien. Auch eine bessere Aufklärung über die im URZ vorhandenen und geplanten Angebote wurde gewünscht. Insbesondere für ältere Mitarbeiter werden Beratungsangebote als essentiell angesehen. Außerdem wurde von einem Befragten der Vorschlag gemacht, dass Mitarbeiter der KS FDM z.B. zwei Stunden in einer Veranstaltung der grundständigen Lehre übernehmen.

Es herrschte Einigkeit darüber, dass eine Verpflichtung der Studenten und insbesondere der Mitarbeiter zum FDM wenig erfolgversprechend ist. Es sollten sinnvolle Vorgehensweisen anhand von Beispielen exemplarisch aufgezeigt werden (z.B. auch für bereits nachgenutzte Daten) und anstatt von „Blacklists“ (was sollte alles nicht gemacht werden) positive Formulierungen verwendet werden (**Pilotprojekte, Best Practice**) (drei Befragte). Projektassoziierte Beispiele sollen dazu führen, dass der Mehrwert durch gutes FDM erkannt wird. Besonders hilfreich wäre auch **individuelle Beratung** bei akuten Problemen (problem- und fachspezifisch). Dafür sollte sich die KS FDM einen Einblick in den Alltag der Wissenschaftler verschaffen und ein tieferes Verständnis für deren Arbeit entwickeln (ein Befragter). Insgesamt betrachtet, ist es eine wichtige Aufgabe, Überzeugungsarbeit zu leisten.

4 Diskussion

Anmerkungen zur Interpretation

Bei allen Interviewpartnern war eine generelle Offenheit gegenüber dem Thema zu spüren. Dies zeigte sich auch darin, dass alle Interviewanfragen positiv beantwortet wurden. Positiv war auch die Bereitschaft zur kritischen Auseinandersetzung mit der eigenen FDM-Praxis bzw. der im Projekt und der offenen Ansprache von Defiziten. Dazu beigetragen hat auch die Zusage, die Interviews nicht im Original und ohne namentliche Zitierung zu veröffentlichen. Jedoch können sich durch die Zusammenfassung der einzelnen Aussagen der Interviews und dem Verzicht auf Personen- und Projektnamen im vorliegenden Bericht geringe Unschärfen ergeben. Für die weitere inhaltliche Bearbeitung im Projekt ist dies nicht relevant, da intern die vollständigen Interviews (Audio/Transkription) vorliegen und verwendet werden.

In der Interviewanfrage an die einzelnen Personen, wurden konkrete Projekte genannt, die Bestandteil des Interviews sein sollten. Dies wurde auch zu Beginn des Interviews noch einmal betont. Die Befragten berichteten aber natürlich aus Ihrem Alltag und ihren allgemeinen Erfahrungen. So kam es häufig vor, dass Bezug auf Vorgängerprojekte genommen wurde, insbesondere, wenn sich das für das Interview ausgewählte Projekt noch in einer Vorbereitungsphase befand. Umgekehrt wurde auch diskutiert, welche Aspekte des FDM in zukünftigen Projekten stärker im Fokus stehen sollten (vor allem aus der Erfahrung mit früheren Projekten). Letztlich wurde sich auch auf andere laufende Projekte bezogen, in denen die Befragten involviert sind. Somit war mitunter nicht klar, ob sich eine Aussage tatsächlich auf das zur Befragung ausgewählte Projekt bezieht.

Genauso war nicht immer eindeutig ersichtlich, inwiefern die getroffenen Aussagen auf das gesamte Verbundprojekt zutreffen. Oft bezogen sich die Befragten auf ihre Institution, auf bestimmte Teilprojekte oder ihre unmittelbare Arbeitsgruppe. Das unterschied sich auch je nachdem, ob ein Sprecher des Verbundprojektes, ein Teilprojektleiter oder ein Datenmanager befragt wurde. Die Sprecher hatten im Allgemeinen einen guten Überblick über das Gesamtbild und die allgemeinen Ziele

und Vorgehensweisen, aber weniger zu Details der praktischen Umsetzung. Die Teilprojektleiter bezogen sich meist auf Ihr Teilprojekt und ihre Arbeitsgruppe und antworteten praxisorientiert. Dagegen konnten die Datenmanager viel zur technischen und organisatorischen Umsetzung und den damit verbundenen Problemen berichten, sich aber nur auf diejenigen Teilprojekte beziehen, mit denen sie regelmäßig im Kontakt standen.

Bei der Auswertung der Interviews wurde außerdem deutlich, dass Begrifflichkeiten bei den Befragten nicht immer klar definiert und einheitlich verwendet wurden. Beispielsweise wurde nicht überall klar, inwiefern die Begriffe „Policy“, „Leitlinie“ oder auch „Richtlinie“ voneinander abgegrenzt werden und welche Verbindlichkeiten sich daraus für das Datenmanagement ergeben. Daher lag stets ein besonderes Augenmerk auf dem Kontext, in dem die jeweiligen Begriffe genannt wurden, um eine möglichst konsistente Auswertung der Ergebnisse zu garantieren.

Insgesamt sollten die individuellen Aussagen der verschiedenen Gesprächspartner nicht als eine repräsentative Stichprobe oder als generalisierbar betrachtet werden. Hierfür ist die Anzahl der durchgeführten Interviews zu gering. Dennoch ergibt sich, nicht zuletzt vor dem eigenen Erfahrungshintergrund der Autoren, ein aufschlussreiches Gesamtbild, das gewisse Schlussfolgerungen zum Stand und Bedarf zulässt und als Grundlage für die strategische Ausrichtung des FDM an der FSU dienen kann.

Stand und Bedarf in den Verbundprojekten an der FSU

Im Vergleich zu den Befragungsergebnissen an anderen Hochschulen (vgl. Tab.1) ergaben die zehn Interviews in den ausgewählten Verbundprojekten an der FSU Jena z.T. ähnliche Aussagen im Hinblick auf grundsätzliche Präferenzen der Befragten (etwa das Bedürfnis Forschungsdaten in der unmittelbaren Umgebung zu halten), allgemeine Entwicklungstendenzen (z.B. wachsendes Bewusstsein für ein nachhaltiges FDM) und fachspezifische Ausprägungen (z.B. restriktiver Umgang mit Forschungsdaten bei industrienaher Forschung).

Bei der Mehrheit der befragten Verbundprojekte wurde sich in der **Planungsphase** mit Aspekten (z.B. „10-Jahre-Regel“ und Langzeitspeicherung) des FDM auseinandergesetzt, jedoch ohne dies detailliert in einem Datenmanagementplan festzuhalten. Das Beratungsangebot der KS FDM existierte zu Beginn der meisten Verbundprojekte noch nicht. Einige der Befragten gaben an, sich bei zukünftigen (Verbund-)Projekten intensiver mit den Aspekten des FDM zu befassen und zeigten sich offen für gezielte Beratungsangebote in der Antragsphase. Dies zeigt, dass bei den Verantwortlichen eine **Sensibilisierung** für das Thema stattgefunden hat, jedoch weiterhin ein erheblicher Informations- und Aufklärungsbedarf hinsichtlich der konkreten Umsetzung der von der DFG herausgegebenen Leitlinien (siehe auch Abschnitt 3.4) zum Umgang mit Forschungsdaten besteht. In diesem Zusammenhang steht auch der Bedarf an **Schulungen** (z.B. differenzierte Angebote für Promovenden und Studierende unterschiedlicher Fachrichtungen nach dem Bachelor-Abschluss) und eine langfristige Integration in die grundständige Lehre. Hierfür wurden durch die Befragung einige Ansatzpunkte für eine konkrete Umsetzung gefunden.

Ein wesentliches Charakteristikum von Verbundprojekten ist das kollaborative Arbeiten in einem **heterogenen und fragmentierten Forschungsumfeld**. Durch die Befragungen wurde deutlich, dass diese Eigenschaft insbesondere für das Thema FDM prägend ist. Oder, um es mit den Worten des Rfil (2016) zu sagen, beim Thema FDM in Verbundprojekten kann „im Positiven von Vielfalt, im Negativen von Zersplitterung“ gesprochen werden. Die Vielfalt zeigt sich z.B. in fachspezifischen Forschungskulturen einzelner Partner, ihren unterschiedlichen Zielsetzungen (Forschung mit oder ohne wirtschaftliche Verwertungsabsicht), der Bedeutung von Forschungsdaten für die jeweilige Forschungsfrage (kaum digitale Daten vs. große Messkampagnen) oder ihrem Bewusstsein und ihrer Bereitschaft FDM im Sinne von Open Science zu betreiben (Nachvollziehbarkeit vs. Nachnutzbarkeit).

Von Zersplitterung kann gesprochen werden, wenn es gilt, die verschiedenen administrativen, organisatorischen und institutionellen Gegebenheiten (z.B. universitär, außeruniversitär, Privatwirtschaft) der beteiligten Partner zu berücksichtigen und sich im Projekt z.B. auf eine gemeinsame Data Policy oder Infrastruktur zu einigen. Herausforderungen dieser Art finden sich auch auf der Ebene einzelner Personen wieder, wenn diese z.B. zu mehreren Organisationen gehören und sich daraus **Konflikte aufgrund unterschiedlicher Leit- und Richtlinien** ergeben.

In den befragten Projekten bestehen z.T. große Unterschiede hinsichtlich des Entwicklungsstandes, der Ausstattung und der verfügbaren Ressourcen für ein FDM (u.a. technische Infrastruktur, Vorbildung der Mitarbeiter, Personalmittel). Eine Zersplitterung zeigt sich insbesondere dadurch, dass in verschiedenen (Verbund-)Projekten für ähnliche Probleme oftmals unterschiedliche Lösungen gefunden wurden. Hier wäre eine **übergreifende Einschätzung** seitens einer zentralen Institution zum Bedarf an und der praktischen Umsetzung von FDM (unter Berücksichtigung der konkreten fachbezogenen Forschungsfragen und individuellen Erfahrungen der Forscher auf diesem Gebiet) von großem Nutzen für die Beteiligten. Durch einen **koordinierenden und vermittelnden Beratungsansatz** (z.B. durch die KS FDM), idealerweise schon frühzeitig während der Planungsphase, könnten einzelne Projekte an der FSU „direkt“ von den Erfahrungen anderer profitieren. Aus Sicht der Projekte selbst wäre diese Art des projektübergreifenden Vorgehens bei der Wahl adäquater Standards und (Spezial-)Lösungen nach Auffassung der Autoren nicht möglich. Eine weitere Möglichkeit hierfür wäre jedoch eine stärkere Gewichtung des FDM über die in DFG Sonderforschungsbereichen beantragbaren Teilprojekten für Informationsinfrastrukturen (INF-Projekt). Obwohl diese in den Gesprächen mehrfach erwähnt wurden und als ein gutes Mittel betrachtet werden um FDM institutionell im Projekt zu verankern sowie mit den notwendigen Ressourcen auszustatten, verfügen nur wenige der befragten SFBs über ein INF-Projekt. Daraus lässt sich ein gewisser **Beratungsbedarf für INF-Projekte** ableiten.

Abschließend ist zu erwähnen, dass das Thema **Langzeitarchivierung** weiterhin viele offene Fragen hinterlässt, die nur im Zusammenspiel aller Beteiligten im Projekt und vor dem Hintergrund der individuellen Anforderungen der Forschenden zufriedenstellend bearbeitet werden können:

1. Welche Daten sollten wie lange aufbewahrt werden?
2. Welche Kriterien können und müssen dafür angesetzt werden?
3. Wer führt die Migrationen von Datenformaten, Datenträgern, etc. durch, damit diese über lange Zeiträume lesbar und für den Forschenden verwertbar bleiben?

Der Bedarf an **zukunftssträchtigen Lösungen**, die in der Praxis robust umgesetzt werden können, ist bei Projekten mit einem erheblichen Datenaufkommen groß und wird in der Beratungstätigkeit der KS FDM eine zentrale Rolle spielen.

5 Ausblick

Im weiteren Fortgang des Projektes soll zunächst den Befragten ein Feedback zu den Ergebnissen der Studie gegeben werden, um Ihre Bereitschaft zu den Interviews zu honorieren und sie an der weiteren Diskussion und Planung konkreter Maßnahmen zu beteiligen.

Neue Fragen, die in weiteren Studien beantwortet werden könnten, sind die Wahrnehmung und Umsetzung der Leitlinien zum Umgang mit Forschungsdaten der DFG von Ende 2015 sowie der Leitlinien der FSU zum FDM, die zu Beginn des Jahres veröffentlicht wurden.

Interessant wäre auch die Erweiterung der Studie auf Einzelprojekte. Gerade hier sind eigene Strukturen zum FDM höchstwahrscheinlich kaum vorhanden und zentrale Strukturen der FSU für ein

effektives und effizientes FDM notwendig. Durch eine Erweiterung auf weitere / alle Hochschulen in Thüringen könnte die FSU durch Ihre bisherigen Erfahrungen Unterstützung leisten. Auch eine Überprüfung beantragter, aber nicht geförderter Projekte (mit und ohne vorherige Beratung zu FDM-Themen) könnte eventuelle Problematiken in FDM-Strategien aufdecken und dadurch präventive Maßnahmen bei zukünftigen Vorhaben ermöglichen.

Danksagung

Das Projekt wird gefördert vom BMBF im Rahmen der Förderrichtlinie zur Erforschung des Managements von Forschungsdaten in ihrem Lebenszyklus an Hochschulen und außeruniversitären Forschungseinrichtungen. Unser ausgesprochener Dank gilt unseren Interviewpartnern, die sich trotz voller Terminkalender die Zeit genommen haben, unsere Fragen zu beantworten und sich auf das Thema eingelassen haben. Wir danken Dr. Daniela Siegel, Dr. Franka Baaske und Christoph Rzymiski für ihre beratende Unterstützung zu Design und Durchführung der Studie. Nathalie Roth, Annegret Gräfe, Annemarie Müller und Bettina Kuntz danken wir ebenfalls für hilfreiche Hinweise sowie für die Transkription der Audioaufnahmen. Ein besonderer Dank gilt Prof. Dr. Sylka Scholz und Dr. Ole Karnatz für die Unterstützung bei der Transkription im Methodenlabor des Instituts für Soziologie der FSU.

Referenzen

Bauer, Bruno; Ferus, Andreas; Gorraiz, Juan; Gründhammer, Veronika; Gumpenberger, Christian; Maly, Nikolaus; Mühlegger, Johannes Michael; Preza, José Luis; Sánchez Solís, Barbara; Schmidt, Nora; Steineder, Christian (2015): Forschende und ihre Daten. Ergebnisse einer österreichweiten Befragung – Report 2015. Version 1.2. DOI: 10.5281/zenodo.32043. Online auch unter:

Deutsche Forschungsgemeinschaft (DFG) (2013): Sicherung guter wissenschaftlicher Praxis, Denkschrift. Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft“. Ergänzte Auflage, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

DFG (2013): Empfehlungen zur Sicherung guter wissenschaftlicher Praxis. Denkschrift. Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft“. Ergänzte Auflage (erste Auflage 1998).

DFG (2015): Leitlinien zum Umgang mit Forschungsdaten. Verabschiedet durch den Senat der DFG am 30. September 2015.

Feldsien-Sudhaus, Inken; Rajski, Beate (2016): Digitale Forschungsdaten für die Zukunft sichern: Umfrage zum Umgang mit Forschungsdaten an der TU Hamburg: Auswertung.

Friedrich-Schiller-Universität Jena (2017): Leitlinie zum Umgang mit Forschungsdaten Friedrich-Schiller-Universität Jena. Am 20.12.2016 durch den Senat beschlossen. Veröffentlicht am 20.03.2017.

Friedrich-Schiller-Universität Jena (2017): Handlungsempfehlung zum Forschungsdatenmanagement an der Friedrich-Schiller-Universität Jena. Am 20.12.2016 durch den Senat beschlossen. Veröffentlicht am 20.03.2017.

Hauck, Reingis; Kaps, Reiko; Krojanski, Hans Georg; Meyer, Anneke; Neumann, Janna; Soßna, Volker (2016): Der Umgang mit Forschungsdaten an der Leibniz Universität Hannover. Auswertung einer Umfrage und ergänzender Interviews 2015/16. Institutionelles Repositorium der Leibniz Universität Hannover.

Herwig, Sebastian; Vogl, Raimund; Rudolph, Dominik (2014): Forschungsdatenmanagement an der WWU. Ergebnisse einer Umfrage zu Status Quo & Entwicklungsperspektiven, Vortrag, 1. Workshop der DINI/nestor AG Forschungsdaten "Institutionelle Forschungsdaten-Policies und strategische Planung des Forschungsdatenmanagements", 02.10.2014, Berlin.

Krähwinkel, Esther (2015): Forschungsdatenmanagement an der Philipps-Universität Marburg. Die Ergebnisse der Umfrage zum Forschungsdatenmanagement im November 2014. Herausgegeben vom Projekt Forschungsdatenmanagement und -archivierung.

Kronenwett, Simone (2016): Forschungsdaten an der Philosophischen Fakultät der Universität zu Köln. Masterarbeit, Studiengang Bibliotheks- und Informationswissenschaft (MA LIS), Fakultät für Informations- und Kommunikationswissenschaften, Technische Hochschule Köln.

Langhanke, Gerald; Stille, Wolfgang (2015): Umgang mit Forschungsdaten an der TUD. Erste Schritte zur Bedarfserhebung und Leitlinienentwicklung, Vortrag, Workshop "Forschungsdatenmanagement" der hessischen Hochschulen, 18.06.2015, Marburg,

Lemaire, Marina; Rommelfanger, Yvonne; Ludwig, Jan; Lürken-Uhl, Alexander; Merkler, Benjamin; Sturm, Peter (2016): Umgang mit Forschungsdaten und deren Archivierung. Bericht zur Online-Bedarfserhebung an der Universität Trier (eSciences Working Papers 02), Online:

Paul-Stüve, Thilo; Rasch, Georg; Lorenz, Sören (2015): Ergebnisse der Umfrage zum Umgang mit digitalen Forschungsdaten an der Christian-Albrechts-Universität zu Kiel (2014) [online]. Zenodo, 2015.

RfII (2016): Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland.

Simukovic, Elena; Kindling, Maxi; Schirnbacher, Peter (2013): Umfrage zum Umgang mit digitalen Forschungsdaten an der Humboldt-Universität zu Berlin. Umfragebericht, Version 1.0. Online verfügbar auf dem edoc-Server der Humboldt-Universität zu Berlin. URN: urn:nbn:de:kobv:11-100213001

Simukovic, Elena; Thiele, Raphael; Struck, Alexander; Kindling, Maxi; Schirnbacher, Peter (2014): Was sind Ihre Forschungsdaten? Interviews mit Wissenschaftlern der Humboldt-Universität zu Berlin. Bericht, Version 1.0. Online verfügbar unter: urn:nbn:de:kobv:11-100224755

Tristram, Frank et al. (2015): Öffentlicher Abschlussbericht von bwFDM-Communities. Wissenschaftliches Datenmanagement an den Universitäten Baden-Württembergs.

Anhang

Interviewleitfaden

1) Würden Sie uns bitte in ein paar kurzen Sätzen sagen, worum es in Ihrem Verbundprojekt geht?

Datentypen, Formate und Größe

2) Welchen Anteil der Forschungsdaten im gesamten Projekt generieren Sie geschätzt in digitaler Form? Dazu zählen Rohdaten und auch verarbeitete Daten. Denken Sie dabei auch Entwürfe, Pläne, Skizzen, Abbildungen, Laborbücher, Feldnotizen etc.

3) Welche Quellen liegen Ihren Forschungsdaten zugrunde? Z.B. Beobachtungen, Experimente, Simulationen, Umfragen, Textdokumente, Bilddokumente, ...

4) Wie kann man sich das vorstellen? Beschreiben Sie in ein paar Sätzen den allgemeinen Verarbeitungsprozess der in Ihrem Projekt anfallenden Forschungsdaten hinsichtlich Erfassung und Analyse.

5) Liegen Ihre Forschungsdaten in Standardformaten oder als fach- / geräte- / softwarespezifische Daten vor?

- bei Antwort Standard: 6a) Welchen Standardformate sind das?

- bei Antwort spezifisch: 6b) Welche Geräte/Software benutzen Sie? 6c) Werden die Daten für die weitere Bearbeitung umgewandelt?

7) Wie häufig verwenden oder generieren Sie Forschungsdaten sensibler oder vertraulicher Natur?

8) Bitte schätzen Sie die Gesamtgröße der Forschungsdaten Ihres Projektes bezogen auf den benötigten Speicherplatz. Z.B. Gigabyte-Bereich

Datenarchivierung und -sicherung

9) Haben Sie sich im Vorfeld des Projektes (während der Planungsphase) bereits mit Forschungsdatenmanagement auseinandergesetzt oder erst während der Projektlaufzeit?

11) Gibt es formale Richtlinien oder Prozeduren, die den Umgang mit Forschungsdaten in Ihrem Projekt genauer bestimmen?

12) Beschreiben Sie in ein paar Sätzen den allgemeinen Verarbeitungsprozess der in Ihrem Projekt anfallenden Forschungsdaten hinsichtlich Bereinigung, Dokumentation und Sicherung?

13) Wo werden die Forschungsdaten während der Verarbeitung in der Projektlaufzeit gespeichert?

14) Werden zwischenzeitlich Sicherungskopien erstellt und wenn ja wie oft?

15) Was passiert mit den von Ihnen generierten Forschungsdaten, wenn das Projekt beendet ist?

- 16) Entspricht das den Kriterien der oben genannten Richtlinien/Vereinbarungen?
- 17) Wo werden die Forschungsdaten für die Langzeitarchivierung gespeichert?
- 18) Werden spezielle Erhaltungsmaßnahmen ergriffen und wenn ja, welche?
- 19) Wer ist verantwortlich für die Speicherung, Sicherung oder Archivierung der Forschungsdaten im Projekt?
- 20) Werden Ihre Forschungsdaten beschrieben / Metadaten erfasst? „Beschreibung“ bedeutet zu erklären, wie die Daten erfasst wurden, was sie bedeuten, wie ihre Struktur aussieht und welche Änderungen und Bearbeitungsschritte zur Bereinigung und Analyse der Daten durchgeführt worden sind.
- 21) Wenn ja, gibt es dafür Standards / einheitliche Vorgaben oder ist die Beschreibung individuell?
- 22) Wer ist verantwortlich für die Beschreibung / Erfassung der Metadaten?
- 23) Welche Eigenschaften (Metadaten) werden erfasst?
- 24) Findet eine automatische Erfassung von Metadaten an Messgeräten / Analysesystemen statt?
- 25) Wenn ja, welche Eigenschaften werden bereits am Gerät erfasst?
- 26) In welcher Form findet bei der Archivierung eine inhaltliche / wissenschaftliche Qualitätssicherung statt?
- 27) In welcher Form findet bei der Archivierung eine technische Qualitätsprüfung statt?

Zugänglichkeit und Nachnutzung

Was ist in Ihrer Fach-Community in Bezug auf Datenpublikationen (falls vorhanden), das Teilen von Daten (data sharing) bzw. die Nutzung von Fremddaten üblich?

- 28) Wie häufig werden Daten ausgetauscht?
- 29) Werden für den Austausch standardisierte Data-Sheets verwendet oder ist eine aufwendige Nachbearbeitung von Fremddaten nötig?
- 30) Wer außer Ihnen darf bzw. kann auf Ihre Forschungsdaten zugreifen?
- 31) Wie können andere auf Ihre Forschungsdaten zugreifen?
- 32) Gibt es Sperrzeiten oder andere Einschränkungen?
- 33) Wird die Auffindbarkeit der Daten durch bestimmte Strategien erhöht und wenn ja, welche?
- 34) Gibt es einen konkreten Plan, die Nachnutzbarkeit der Daten zu fördern und wenn ja, wie?

Lehre

35) Welche Vorkenntnisse im Forschungsdatenmanagement bringen Ihre Projektmitarbeiter (inkl. Doktoranden) mit?

36) Führen Sie Lehrveranstaltungen / Trainings zu Forschungsdatenmanagement durch?

37) Halten Sie Lehrveranstaltungen / Trainings zu Forschungsdatenmanagement für sinnvoll / Würden Sie angebotene Materialien einsetzen?

IT-Infrastruktur und Service

38) Nutzen Sie Managementtools in der Projektphase?

39a) Wenn nein, warum (nicht)?

39b) Wenn ja, welche?

40) Sehen Sie Bedarf für solche Tools / würden Sie diese nutzen?

41) Von welchen Infrastrukturangeboten der FSU für Forschungsdaten haben Sie bereits gehört

42) Welche davon nutzen Sie bereits?

43) Nutzen Sie die technische Infrastruktur anderer Einrichtungen (z.B. in gemeinsamen Projekten)?

44) Welche Serviceleistungen in Bezug auf Forschungsdatenmanagement würden Sie sich von der FSU wünschen?

45) Gibt es sonst noch Punkte in Bezug auf Forschungsdatenmanagement, in denen Sie Verbesserungsbedarf an der FSU sehen?

46) Haben Sie weitere Hinweise und Anmerkungen?