Forschungsdatenmanagement Bayern

UMGANG MIT DIGITALEN FORSCHUNGSDATEN IN DEN GEISTES- UND SOZIALWISSENSCHAFTEN

Bericht zur Bedarfserhebung an der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

FDM-BAYERN.ORG

Autorin: Anna Daudrich

September 2018

INHALTSVERZEICHNIS

1. Einleitung **3**

2. Methode **4**

2.1 Online-Umfrage 4

2.2 Leitfaden-Interviews 5

3. Ergebnisse **6**

3.1 Fach- und positionsspezifischer Kontext der Teilnehmer **6**

3.1.1 Fachspezifische Zuordnung 7

3.1.2 Positionsspezifische Zuordnung 8

3.2 Praxisanalyse **8**

3.2.1 Planung der Datenerhebung 9

3.2.2 Datensammlung und –erzeugung 9

3.2.2.1 Art der anfallenden Forschungsdaten 9

3.2.2.2 Quellen der Forschungsdaten 11

3.2.2.3 Umfang der anfallenden Daten 11

3.2.3 Beschreibung im Entstehungskontext 13

3.2.4. Aufbereitung der Daten 14

3.2.4.1 Dateiformate 14

3.2.4.2 Datenbeschreibung 16

3.2.5 Datenspeicherung und -sicherung 17

3.2.5.1 Speicherorte 17

3.2.5.2 Backup 18

3.2.6 Datenaufbewahrung 19

3.2.7 Zugänglichkeit und Nachnutzung 19

3.3 Anforderungs- und Bedarfsanalyse **21**

3.3.1 Anforderungen an ein Datenmanagementplan-Tool 21

3.3.2 Bedarf nach Veröffentlichung bzw. Zugänglichmachung 22

3.3.3 Beratungs- und Unterstützungsbedarf 25

4. Fazit **26**

1. EINLEITUNG

Mit der zunehmenden Verwendung digitaler Methoden in den Geistes- und Sozialwissenschaften steigt auch dort die Menge und Heterogenität der Daten, die in wissenschaftlichen Forschungsprozessen entstehen, rasant an. Das Spektrum reicht von Digitalisaten von Handschriften über Textkorpora und digitale Aufnahmen von Kunstwerken und Artefakten bis hin zu digitalen Aufzeichnungen von Tönen oder Sprachmustern. Die Heterogenität und die Komplexität der Forschungsdaten stellen technische, rechtliche und ethische Herausforderungen dar. Vor dem Hintergrund dieser Entwicklung werden seitens einzelner Hochschulen aber auch institutionsübergreifend zahlreiche Initiativen eingeleitet und Services aufgebaut, die das effiziente Management von Forschungsdaten mit Blick auf ihre Erfassung, Aufbereitung, Dokumentation, Archivierung und Bereitstellung unterstützen und fördern.

Die Universitätsbibliothek der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) widmet sich ebenfalls verstärkt dieser Aufgabe. In diesem Zusammenhang wurden von Mai bis Juli 2018 alle Wissenschaftler[1] der Philosophischen Fakultät und dem Fachbereich Theologie der Universität zum Umgang mit ihren Forschungsdaten befragt. Diese Befragung verfolgte das Ziel, die gegenwärtige Situation des Forschungsdatenmanagements (FDM) in den digitalen Geistes- und Sozialwissenschaften an der FAU zu eruieren und den Bedarf an entsprechender Schulung und Beratung zu ermitteln.

Die Studie ist Teil des Modellvorhabens „eHumanities – interdisziplinär“, das vom Bayerischen Staatsministerium für Wissenschaft und Kunst im Rahmen des „Digitalen Campus Bayern“ für eine Laufzeit von drei Jahren (2018 - 2020) gefördert wird. Mit Fokus auf das interdisziplinäre Fachgebiet der digitalen Geistes- und Sozialwissenschaften werden seitens der Universitätsbibliothek der FAU gemeinsam mit der Universitätsbibliothek der Ludwig-Maximilians-Universität München (UB LMU) und der IT-Gruppe Geisteswissenschaften der LMU (ITG) relevante Tools, Infrastrukturen sowie Best-Practice-Empfehlungen für das FDM erarbeitet.

Ein weiteres Ziel des Projekts ist die Unterstützung der Lehre und Fortbildung bei der Vermittlung von FDM-Kompetenzen. Hierfür wird ein auf die Bedürfnisse der interdisziplinären Geistes- und Sozialwissenschaften angelegtes zeit- und ortsunabhängiges E-Learning-Angebot entwickelt. Dieses soll auf modularer Basis realisiert werden und sowohl fachübergreifende als auch fachspezifische Themen beinhalten. Im Probelauf werden die entwickelten Module zunächst an der FAU und im Anschluß an die darauffolgende Evaluation und Verbesserung auch an der LMU eingesetzt. Eine breite, institutionsübergreifende Nachnutzbarkeit der Lehrangebote als Open Educational Resources ist dabei ein zentrales Ziel.

[1] Aus Gründen der besseren Lesbarkeit wird im Rahmen der vorliegenden Arbeit ausschließlich die männliche Form verwendet. Sämtliche Personenbezeichnungen gelten gleichermaßen für alle Geschlechter.

Die durchgeführte Bedarfserhebung wird als Ausgangsbasis für die Konzeption der Module dienen. Darüber hinaus werden die Ergebnisse der Studie in die Weiterentwicklung der Beratungsangebote zum FDM für Graduierte und Studierende der FAU einfließen.

2. METHODE

Die Studie basiert auf einer Online-Umfrage sowie ergänzenden Leitfaden-Interviews und wurde vom Referat Forschungsdatenmanagement der Universitätsbibliothek der FAU durchgeführt.

2.1 ONLINE-UMFRAGE

Ziel der Online-Umfrage war die Gewinnung erster Erkenntnisse über den Umgang mit digitalen Forschungsdaten in den geistes- und sozialwissenschaftlichen Forschungsbereichen an der FAU.

Die Umfrage richtete sich an wissenschaftliche Angehörige der Philosophischen Fakultät und des Fachbereichs Theologie. Die Zielgruppe wurde über einen E-Mail-Verteiler kontaktiert, der 2191 Mitglieder umfasste. Die Befragten hatten zudem die Möglichkeit, die Umfrage an Masterstudierende ihrer Wahl weiterzuleiten. Bachelorstudierende wurden von der Befragung ausgeschlossen, da Forschungsdaten in dieser Phase der akademischen Laufbahn nur bedingt erhoben werden. Die Teilnahme an der Umfrage war über drei Wochen hinweg möglich. Ausgewählte Wissenschaftler wurden zudem mit der Bitte um die Weitergabe der Umfrage an Forschende ihres Instituts bzw. Lehrstuhls persönlich kontaktiert.

Insgesamt nahmen an der Umfrage 86 Personen teil. Da sich die Befragung ausschließlich auf die Arbeit mit digitalen Forschungsdaten bezog, sollten die Teilnehmer gleich zu Beginn angeben, ob digitale Forschungsdaten im Rahmen ihrer Forschungstätigkeit relevant sind. Im Fall einer verneinenden Antwort wurde die Umfrage beendet, da der Teilnehmer nicht der untersuchten Zielgruppe entsprach. 70 Befragte gaben an, digitale Forschungsdaten in ihre wissenschaftliche Tätigkeit einzubeziehen. Diese Forschende konnten die Beantwortung der Fragen fortsetzen. Der Fragebogen wurde dabei in drei Abschnitte unterteilt und orientierte sich wesentlich an den bereits durchgeführten Befragungen anderer wissenschaftlicher Institutionen[2].

Zunächst sollten die Befragten ihre Zugehörigkeit zu einem bestimmten Fach- bzw. Forschungsbereich sowie ihre Position im jeweiligen Bereich angeben. Diese Informationen dienen der Formulierung von fach- und karrierespezifischen Aussagen. Die identifizierten disziplinspezifischen Gemeinsamkeiten und Unterschiede werden bei der Planung und Konzeption fachspezifischer Lerninhalte des E-Learnings berücksichtigt. Des Weiteren erlauben diese Erkenntnisse fachspezifisch angemessene Konzepte und Überlegungen für die Qualitätssicherung der Forschungsdaten zu entwickeln. Da sich Lehr- und Beratungsangebote zum FDM an der FAU an verschiedene

[2] http://www.forschungsdaten.org/index.php/Umfragen_zum_Umgang_mit_Forschungsdaten_an_wissenschaftlichen_Institutionen (Stand: September 2018).

Zielgruppen (z.B. Studierende und Doktoranden) richten, war auch die Frage nach der aktuellen Position unerlässlich. Infolgedessen werden karrierespezifische Anforderungen im Rahmen der Studie gleichermaßen untersucht. Fach- und karrierespezifische Differenzierungen belegen zudem, ob förderliche Kompetenzen und die Sensibilität für die Thematik Forschungsdaten in allen Gruppen ausreichend vorhanden sind.

Der zweite Teil des Fragebogens diente im Wesentlichen der Ermittlung eines Überblicks über die aktuelle Art und Weise des Umgangs mit Forschungsdaten. Hierfür wurden Fragen zum Umgang mit Forschungsdaten aus der Sicht eines konkreten geistes- oder sozialwissenschaftlichen Projektes beantwortet. Unter Projekten wurden dabei alle genehmigten Wissenschaftsprojekte verstanden, die unter anderem Abschlussarbeiten, Dissertationen oder auch Gruppenprojekte umfassen. Da das FDM je nach Forschungsvorhaben bzw. Einzelprojekt abweichen kann, wurde durch diese Spezifizierung angestrebt, ein möglichst konkretes Bild zum Status Quo zu erhalten.

Anschließend wurde der allgemeine Umgang mit sowie die Nutzung von Forschungsdaten erfragt. Dieser Teil des Fragebogens diente primär der Bestimmung von Wünschen und Bedürfnissen der untersuchten Zielgruppe.

2.2 Leitfaden-Interviews

Im Anschluss an die Online-Umfrage erfolgte eine zweite Phase der Erhebung, die anhand von Leitfaden-Interviews durchgeführt wurde. Ziel der Leitfadengespräche war es, die aus der Online-Umfrage gewonnenen Erkenntnisse zu bestätigen, zu vertiefen oder gegebenenfalls auch zu modifizieren. Darüber hinaus wurde während der Interviews auf Fragen eingegangen, die im Rahmen der vorangegangenen Online-Umfrage nicht beantwortet werden konnten bzw. hinsichtlich der Aussagekraft und Interpretierbarkeit Schwachstellen aufwiesen.

Zur Teilnahme an einem Interview wurden Professoren, Privatdozenten sowie Promovierende der Fakultät eingeladen, wobei Wissenschaftler aus möglichst unterschiedlichen geistes- und sozialwissenschaftlichen Disziplinen berücksichtigt wurden. Im Rahmen der zur Verfügung stehenden Ressourcen konnten insgesamt neun Personen aus folgenden Forschungsbereichen befragt werden: Altertums- und Geschichtswissenschaften, Buchwissenschaft, Digital Humanities, Germanistik, Kunstgeschichte, Romanistik, Pädagogik.

Im Vorfeld jedes Interviews erhielten die Befragten einen Leitfaden mit vordefinierten Fragen und Themenkomplexen. Die Festlegung der Themenkomplexe orientierte sich an üblichen Phasen eines Forschungsprojektes: Planung und Vorbereitung, Durchführung sowie Abschluss und Nachbereitung. Hierfür wurde beschrieben, welche Aktivitäten, Entscheidungen und Kompetenzen in jede dieser Phasen aus Sicht des FDMs förderlich sind. In den Interviews sollten die Befragten den typischen Umgang mit Forschungsdaten in ihren Fachbereichen erläutern, indem sie sich an diesem Leitfaden orientierten.

Des Weiteren wurde im Rahmen von Interviews auf die Rolle einzelner Lehrstühle und der Universitätsbibliothek in Fragen der Kompetenzvermittlung eingegangen. In diesem Zusammenhang wurde unter anderem diskutiert, in welcher Phase der akademischen Laufbahn (z.B. Bachelor-, Masterstudium, Promotion oder Post-Doc) welche FDM-Kompetenzen (z.B. Kenntnisse in Urheberrecht und über Dateiformate) und in welchem Rahmen (z.B. Schulungsangebote der Universitätsbibliothek vs. Bestandteil des Curriculums) vermittelt werden sollten.

3. Ergebnisse

Im folgenden Abschnitt werden zentrale Ergebnisse der Studie vorgestellt.

3.1 Fach- und positionsspezifischer Kontext der Teilnehmer

Eine Besonderheit der Philosophischen Fakultät der FAU ist, dass sie neben geisteswissenschaftlichen Disziplinen auch die Bereiche aus Wirtschaft- und Sozialwissenschaften sowie Erziehungswissenschaften und Theologie abdeckt. Die Fakultät gliedert sich in zwölf Departments, wobei einige Departments mehrere Institute umfassen, die wiederum einzelne Lehrstühle beherbergen. Insgesamt sind an der Fakultät über 50 verschiedene Disziplinen vertreten, wobei 14 interdisziplinäre Zentren dort zusätzlich angesiedelt sind.

Aufgrund der Vielfalt der vertretenen Disziplinen und Forschungsbereiche wurde im Rahmen der Online-Umfrage zunächst vorgesehen, die Antwortoptionen hinsichtlich der fachspezifischen Zuordnung möglichst differenziert darzustellen. Die Problematik einer präzisen Zuteilung bestand jedoch darin, dass die Angaben zum fachspezifischen Kontext in Kombination mit der Antwort zur Position Rückschlüsse auf die Identität einzelner Datengeber zulassen würden.

Um Anonymität der Teilnehmer zu gewährleisten, wurde beschlossen, kleinere Disziplinen zusammengefasst oder gar nicht eigenständig aufzuführen. In der finalen Version des Fragebogens folgten somit die festgelegten Antwortoptionen grob den Departments der Philosophischen Fakultät; nur in Fällen, in denen ein Department sehr unterschiedliche oder große Disziplinen umfasst, wurde dies aufgeteilt (s. Abb. 1). Auch die seltener vertretenen Positionen wie etwa Privatdozent, Postdoktorand oder Akademischer Rat wurden zur Wahrung der Anonymität zusammen mit anderen vergleichbaren Gruppen in Antwortoptionen gebündelt (s. Abb. 2).

3.1.1 Fachspezifische Zuordnung

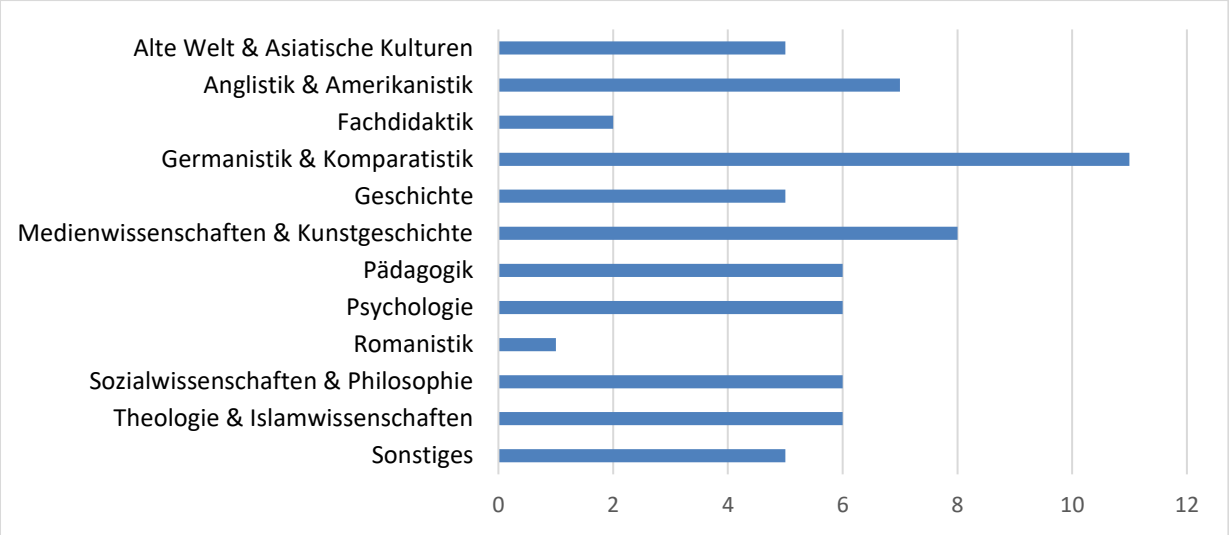

Abb. 1: Beteiligung an der Online-Umfrage nach Forschungsbereichen

Wie aus den Antworten hervorgeht (s. Abb. 1), zeigt der Anteil der Teilnehmer an der Online-Umfrage eine relativ gleichmäßige Verteilung auf die Fachbereiche. Lediglich Romanistik und Fachdidaktik sind deutlich unterrepräsentiert, während die Teilnehmer aus der Germanistik und Komparatistik etwas hervortreten. Eine verhältnismäßig starke Beteiligung der Forschenden aus der Germanistik und Komparatistik ist womöglich darauf zurückzuführen, dass eine der Schwerpunktsetzungen des gleichnamigen Departements in den Themenfeldern der technisch orientierten Korpus- und Computerlinguistik liegt, in deren Rahmen eine besondere Sensibilität für die FDM-Thematik gefragt ist. Überraschend ist dagegen eine niedrige Teilnahme aus der Romanistik, in der der Einsatz digitaler Forschungsmethoden bekanntlich verbreitet ist.

Diese fachspezifische Zuteilung lässt sich jedoch relativieren. Die Leitfadengespräche bestätigten die Annahme, dass eine große Vielfalt thematischer Schwerpunkte sogar innerhalb einzelner Disziplinen besteht. In diesem Zusammenhang ist es fraglich, ob zum Teil deutlich unterschiedliche Herangehensweisen in der Literaturwissenschaft und Linguistik innerhalb der Philologien eine eindeutige Ermittlung des jeweiligen Forschungskontextes anhand der Angaben “Anglistik & Amerikanistik”, “Germanistik & Komparatistik” und “Romanistik” zulassen. Wie die Interviewteilnehmer betonten, werden unterschiedliche Forschungsansätze oft sogar innerhalb einzelner Teilgebiete herangezogen, sodass an die verarbeiteten Daten unterschiedliche Anforderungen gestellt werden. Beispielhaft hierfür sind die Literaturwissenschaften, die zwar als primär hermeneutisch arbeitende Wissenschaften mit einer überschaubaren Menge anfallender Forschungsdaten gelten; es werden jedoch gelegentlich auch dort empirische Ansätze - wie etwa Distant Reading - zur Verarbeitung großer Datenmengen eingesetzt. Ein ähnliches Bild zeigt sich unter anderem in der Buchwissenschaft, in der sich sowohl historische als auch sozialwissenschaftliche Schwerpunkte finden lassen, die aus Sicht des Forschungsdatenmanagements wenig gemein haben.

3.1.2 Positionsspezifische Zuordnung

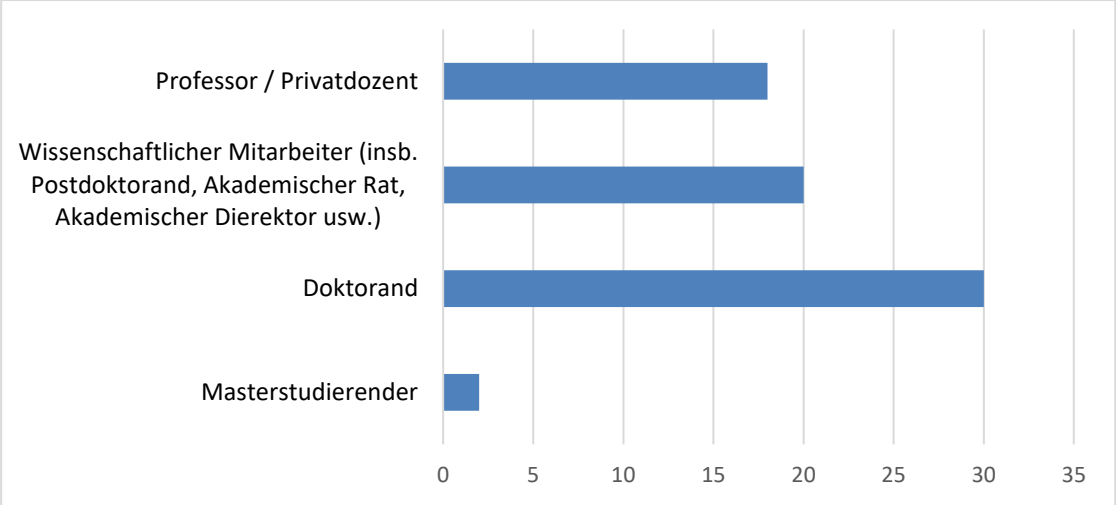

Abb. 2: Beteiligung an der Online-Umfrage nach Karrierestufen

Die Antworten zur Position der Teilnehmer (s. Abb. 2) zeigen, dass die Gruppe "Doktoranden" am stärksten vertreten ist, gefolgt von wissenschaftlichen Mitarbeitern und Professoren bzw. Privatdozenten, wobei sich nur eine geringe Zahl der Masterstudierenden an der Befragung beteiligte. Angesichts einer hohen Gesamtzahl aller Doktoranden ist die Teilnahme dieser Gruppe jedoch als zurückhaltend zu bewerten. Die Beteiligung der Professoren und Privatdozenten ist dagegen als repräsentativer zu betrachten, da diese Positionen im Vergleich zu den anderen genannten Gruppen an der Fakultät in deutlich geringerer Zahl vorhanden sind. Eine hohe Teilnahmequote der Professoren und Privatdozenten lässt die Annahme zu, dass die Sensibilität für die FDM-Thematik in dieser Gruppe am stärksten ausgeprägt ist, was jedoch angesichts ihrer Erfahrung in der Arbeit mit wissenschaftlichen Daten auch zu erwarten war.

Ähnlich wie im Fall einer fachspezifischen Zuordnung lassen sich die Aussagen hinsichtlich der Position innerhalb des jeweiligen Forschungskontextes ebenfalls relativieren. So ist etwa ein Doktorand, der gleichzeitig eine Tätigkeit als wissenschaftlicher Mitarbeiter an einem Lehrstuhl ausübt, in der Liste der vorgegebenen Antwortoptionen nicht eindeutig zuzuordnen. Insofern sind Angaben solcher Teilnehmer vermutlich sowohl in der Gruppe "Doktoranden", zu der sich höchstwahrscheinlich Promovierende ohne Anstellungsverhältnis zugeordnet haben, als auch in der Gruppe "wissenschaftlicher Mitarbeiter" zusammen mit den Antworten von Postdoktoranden oder Akademischen Räten zu finden.

3.2 Praxisanalyse

Im folgenden Abschnitt wird beschrieben, wie mit digitalen Daten im Rahmen eines Forschungsprojektes in den Geistes- und Sozialwissenschaften an der FAU typischerweise umgegangen wird.

3.2.1 Planung der Datenerhebung

Bereits vor dem Beginn eines Forschungsvorhabens ist es hilfreich, die Maßnahmen zu planen, die den Umgang mit anfallenden Forschungsdaten in den kommenden Arbeitsphasen bestimmen werden. Für eine sinnvolle Planung der Datenaufbereitung etablierte sich die Erstellung von Datenmanagementplänen (DMP), die ein systematisches, schriftliches Festhalten der wichtigsten Punkte vorsehen. Immer häufiger werden solche Datenmanagementpläne bei der Projektbeantragung durch Forschungsförderer (z.B. DFG) explizit gefordert.

Wie die Leitfaden-Gespräche jedoch ergaben, wird in den Geistes- und Sozialwissenschaften an der FAU in der Regel kein DMP angefertigt. In der Online-Umfrage gaben nur 15% der Teilnehmer an, jemals einen DMP erstellt zu haben (s. Abb. 3).

Abb. 3: Erstellung eines DMPs (Gesamtwertung)

In den Interviews wurde deutlich, dass diese Arbeit häufig aus Zeitgründen und aufgrund der fehlenden Ressourcen nicht vorgenommen wird. Größtenteils werden die Planungsmaßnahmen sogar als wenig notwendig erachtet. In der Vorbereitungsphase eines Forschungsvorhabens werden üblicherweise lediglich grobe Überlegungen getroffen. Nur in Einzelfällen ist von einem durchdachten Ablaufplan die Rede, wobei es sich dabei größtenteils um mündliche und weniger um schriftliche Vereinbarungen handelt.

3.2.2 Datensammlung und –erzeugung

3.2.2.1 Art der anfallenden Forschungsdaten

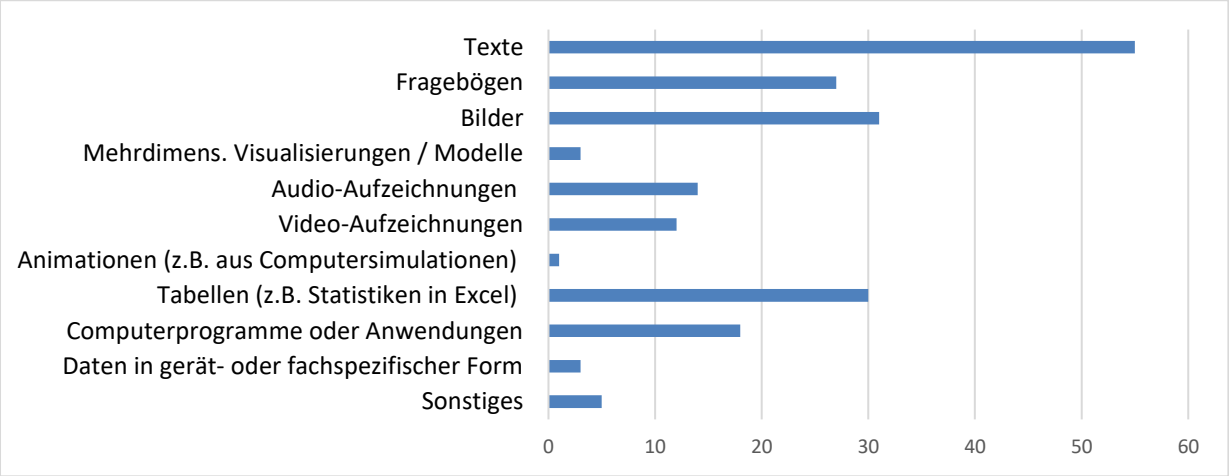

Abb. 4: Art der anfallenden Forschungsdaten (Gesamtwertung; Mehrfachnennung möglich)

Die Landschaft der anfallenden Forschungsdaten ist in der geistes- und sozialwissenschaftlichen Forschungspraxis an der FAU heterogen. Wie die Grafik zeigt (s. Abb. 4), machen den größten Anteil dabei Texte aus. Diesen folgen in größeren Abständen Bilder, Tabellen, Fragebögen sowie Audio- und Videoaufzeichnungen. Auch wenn mehrdimensionale Visualisierungen und Modelle seltener eingesetzt werden, sollen diese Daten dennoch nicht unbeachtet bleiben, da sie in der Regeln ein großes Datenvolumen voraussetzen.

Die Auswertung nach einzelnen Forschungsbereichen ergab (s. Abb. 5), dass der hohe Anteil von Texten durch ihre Relevanz für alle Wissenschaftsdisziplinen zu erklären ist. Andere Forschungsdaten sind hingegen nur in bestimmten Bereichen etabliert. So wird beispielsweise in den Religionswissenschaften und in den Philologien überwiegend mit Texten gearbeitet. Eine deutlich höhere Relevanz der Bilder zeigt sich demgegenüber in Forschungsbereichen wie Alte Welt und Asiatische Kulturen, Geschichte sowie Medienwissenschaften und Kunstgeschichte. Andere Forschungsdaten wie Fragebögen gewinnen dabei in Disziplinen wie Pädagogik und Psychologie deutlich an Bedeutung.

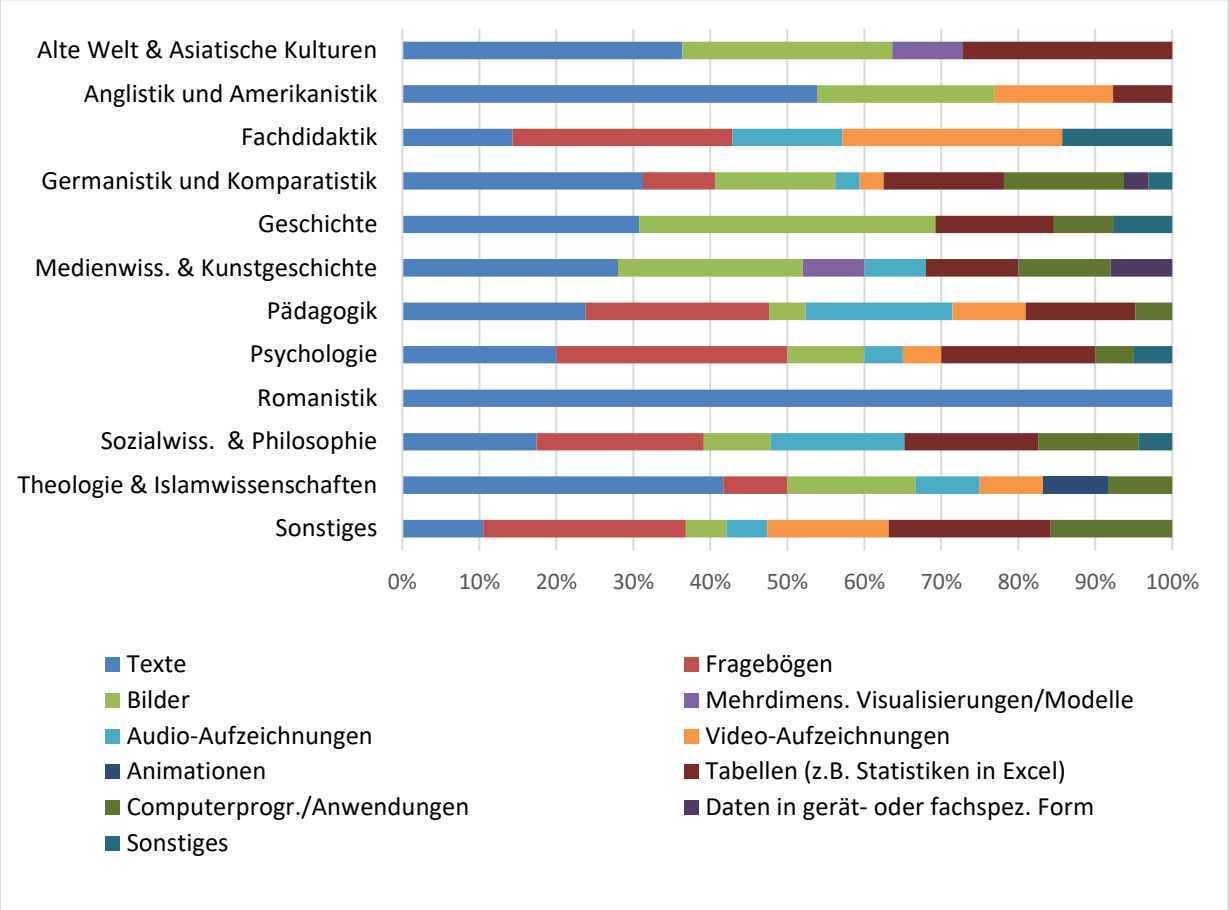

Abb. 5: Art der anfallenden Forschungsdaten (Auswertung nach Fachbereichen)

3.2.2.2 Quellen der Forschungsdaten

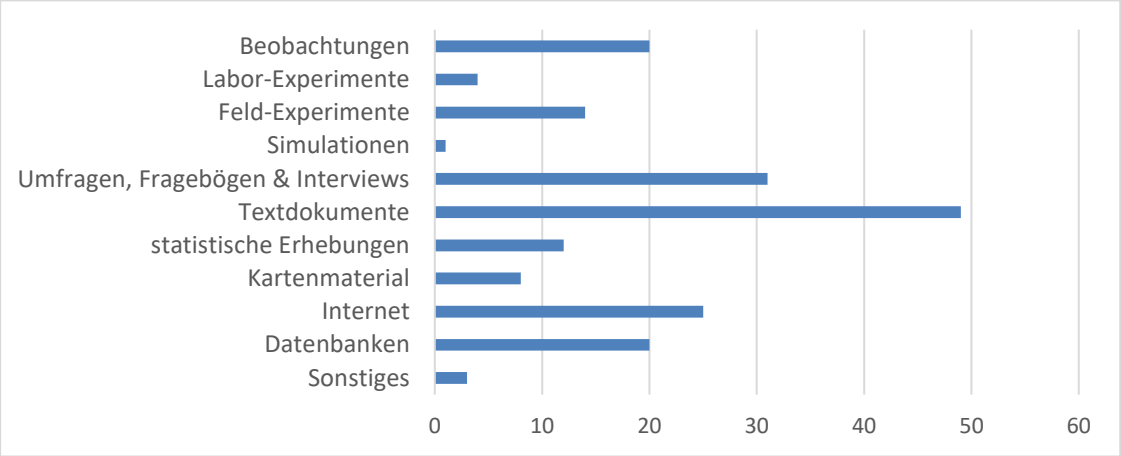

Abb. 6: Quellen der Forschungsdaten (Gesamtauswertung; Mehrfachnennung möglich)

Entsprechend der hohen Heterogenität der anfallenden Forschungsdaten werden auch unterschiedliche Methoden zur Datenerhebung angewendet, wobei Textdokumente erneut die häufigste Nennung aufweisen (s. Abb. 6). Als relevante Datenquellen wurden unter anderem Beobachtungen, Internet und Datenbanken genannt. Da sich innerhalb einiger Fachbereiche der Fakultät sozialwissenschaftliche Schwerpunkte finden lassen, wird ein großer Anteil der Forschungsdaten erwartungsgemäß durch Umfragen, Interviews sowie statistische Erhebungen gewonnen. Simulationen und Labor-Experimente wurden dagegen nur selten genannt, da sie in der Regel mehrheitlich in den technisch-naturwissenschaftlichen Fächern durchgeführt werden.

3.2.2.3 Umfang der anfallenden Daten

Hinsichtlich des Umfangs der anfallenden Forschungsdaten liegen ebenfalls unterschiedliche Angaben vor. Wie die Grafik zeigt (s. Abb. 7), benötigt knapp die Hälfte der befragten Forscher (49%) für ihre wissenschaftlichen Daten weniger als 100 GB, was einem kleinen bis mittleren Speichervolumen entspricht. 9% benötigen bis zu einem TB und 10% weit darüber hinaus. Bemerkenswert ist dabei, dass ein Drittel der Befragten die Speicherkapazität ihrer Forschungsdaten nicht einschätzen konnte.

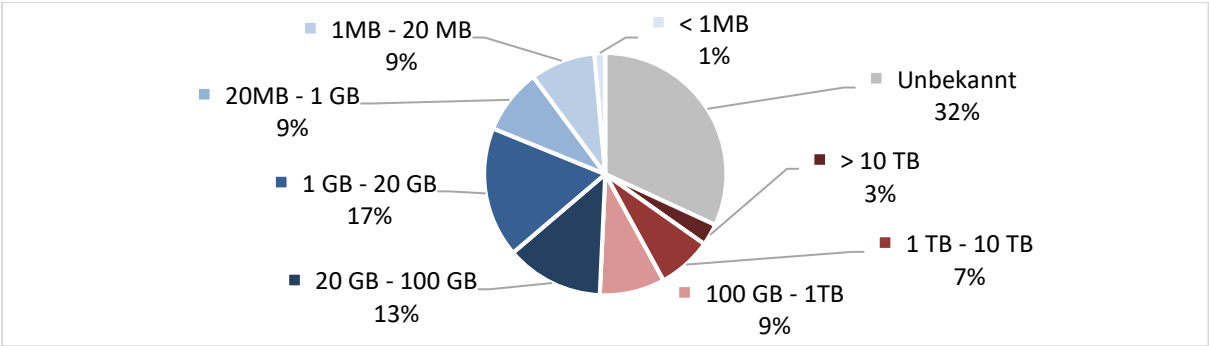

Abb. 7: Benötigter Speicherplatz (Gesamtauswertung)

In Bezug auf die Frage nach der Menge der anfallenden Daten (s. Abb. 8) gaben 61% der Befragten an, 1000 und weniger einzelne Dateien im Rahmen ihrer Forschung zu erzeugen. 22% schätzen die Gesamtgröße ihrer Daten auf 10000 und mehr. Dagegen konnten 17% keine Auskünfte über die Anzahl geben.

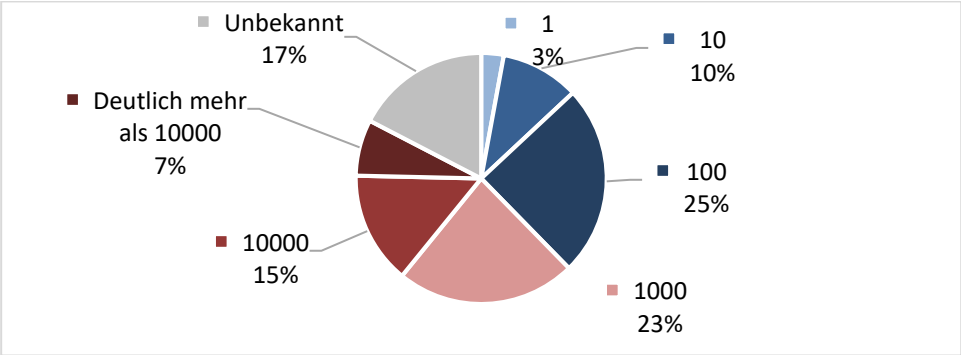

Abb. 8: Anzahl der anfallenden Einzeldateien (Gesamtauswertung)

Die fachspezifische Vergleichsanalyse zeigt, dass in den Wissenschaftskontexten mit einem hohen Anteil an Bildern eine größere Datenmenge und -volumen anfällt, als dies bei textbasierter Forschung der Fall ist (s. Abb. 9, 10). Diese Tendenz wurde im Zuge der Interviews ebenfalls bestätigt.

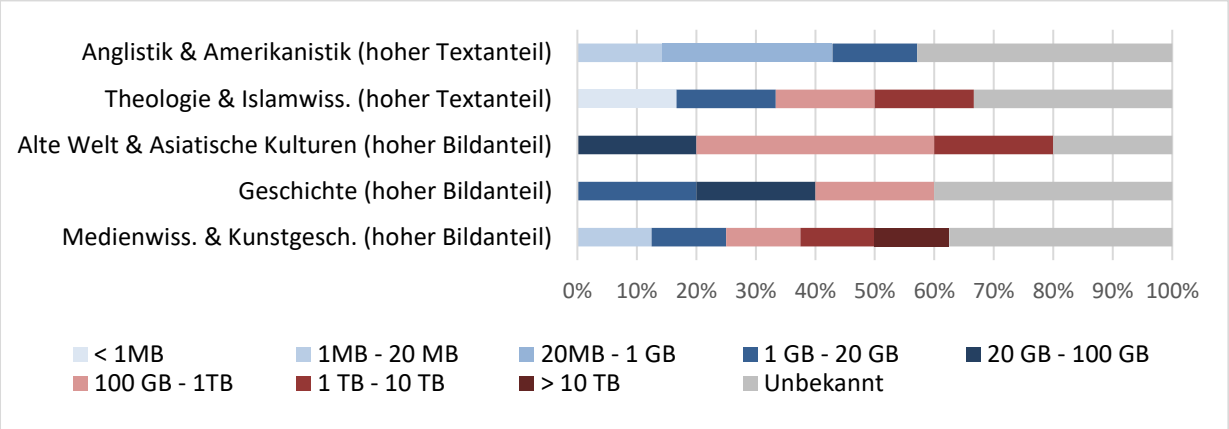

Abb. 9: Benötigter Speicherplatz (Auswertung für Disziplinen mit einem hohen Anteil an Texten bzw. Bildern)

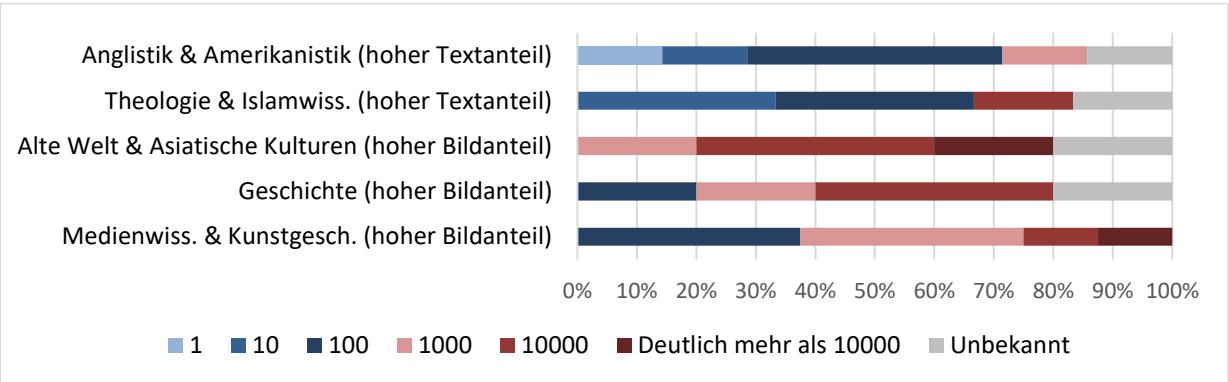

Abb. 10: Anzahl der anfallenden Einzeldateien (Auswertung für Disziplinen mit einem hohen Text.- bzw. Bildanteil)

Als besonders datenintensiv zeigen sich Drittmittelprojekte, unabhängig von der Fachdisziplin (s. Abb. 11, 12). Diese Entwicklung war insofern zu erwarten, als die durch Forschungsförderer finanzierten Forschungsvorhaben in der Regel eine deutlich höhere Komplexität im Vergleich zu meist kleineren universitätsinternen Projekten aufweisen.

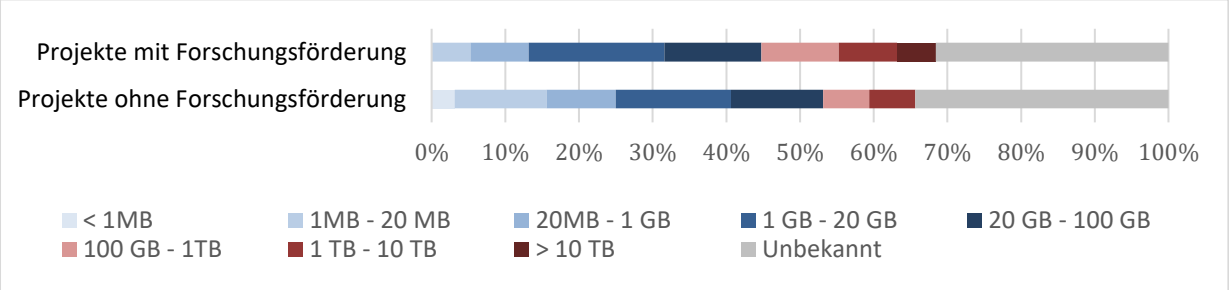

Abb. 11: Benötigter Speicherplatz (Auswertung für Projekte mit und ohne Forschungsförderung)

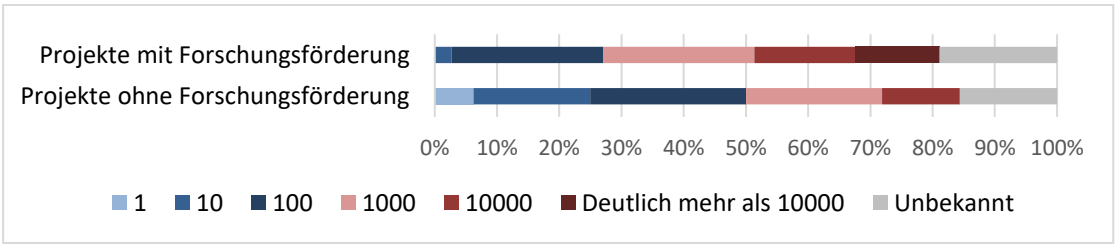

Abb. 12: Anzahl der anfallenden Einzeldateien (Auswertung für Projekte mit und ohne Forschungsförderung)

3.2.3 Beschreibung im Entstehungskontext

Es ist ausgesprochen hilfreich, den Daten-Erhebungsprozess mit beschreibenden Informationen zu dokumentieren, sodass alle wichtigen Schritte nachvollziehbar bleiben. Dies erleichtert die Wiederauffindbarkeit und Rekonstruierbarkeit der Daten bei Verlust oder Leseproblemen.

Wie die Ergebnisse der Online-Umfrage jedoch zeigen (s. Abb. 13), besteht hinsichtlich einer ausführlichen Datendokumentation ein Mangel an Sensibilisierung. Lediglich 46% gaben an, eine Dokumentation für ihre Forschungsdaten anzulegen.

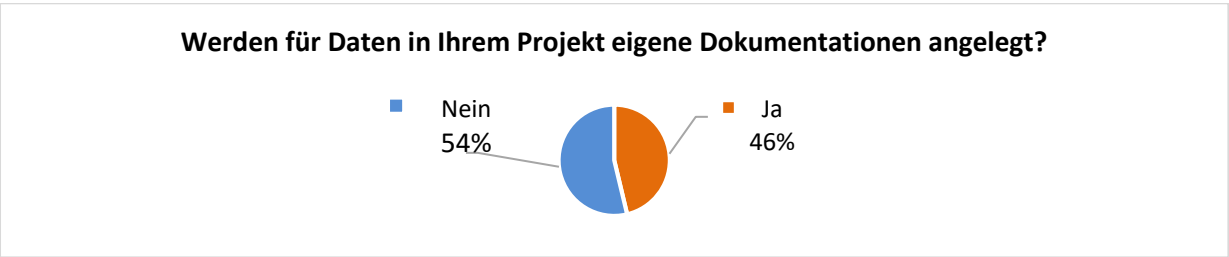

Abb. 13: Datendokumentation (Gesamtauswertung)

Wie aus den Interviews hervorging, wird eine teils zeitaufwändige Dokumentation der Daten häufig als überflüssig und lästig empfunden. Nur in Forschungskontexten mit besonders hohem Datenumfang ist das Bewusstsein hierfür geschärft, da die Handhabung von großen Da-

tenmengen im Falle einer fehlenden Dokumentation schnell zu einem unübersichtlichen Vorgang werden kann. Wie erwähnt, gelten als datenintensiv insbesondere Drittmittelprojekte, in deren Rahmen einer hinreichenden Dokumentation der Forschungsdaten ein verhältnismäßig großer Stellenwert zukommt (s. Abb. 14).

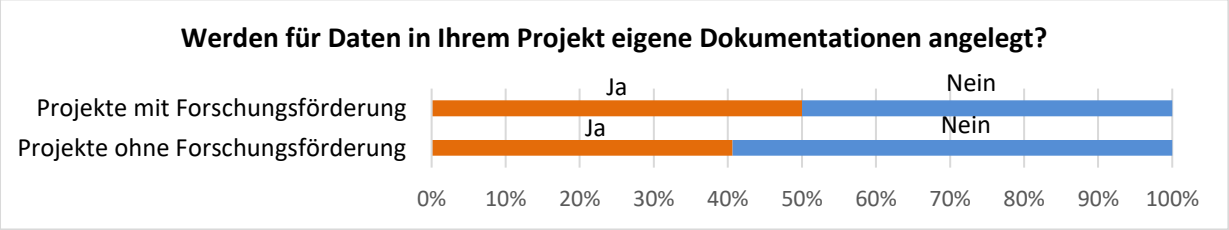

Abb. 14: Datendokumentation (Auswertung für Projekte mit und ohne Forschungsförderung)

3.2.4. Aufbereitung der Daten

Eine effiziente Aufbereitung der Forschungsdaten ist mit Blick auf die Nachhaltigkeit und langfristige Nutzung notwendig. Gut durchdachte Metadaten und Dateiformate spielen in diesem Zusammenhang eine wichtige Rolle.

3.2.4.1 Dateiformate

Bereits bei der Auswahl einer Software sollte berücksichtigt werden, dass veraltete oder proprietäre Dateiformate Hindernisse für die langfristige Nutzung der Forschungsdaten bereiten. Wie die Ergebnisse der Online-Umfrage belegen (s. Abb. 15), zeichnen sich die an der Philosophischen Fakultät verwendeten Dateiformate durch eine große Vielfalt aus, wobei PDF, doc(x), xls(x) gefolgt von tiff/tiff, jpeg/jpg und HTML andere Dateiformate deutlich überwiegen. Auffallend ist, dass proprietäre, lizenzpflichtige Formate wie etwa Microsoft-Office-Formate eine große Verbreitung finden.

Abb. 15: Wortwolke zu verwendeten Dateiformaten (Gesamtauswertung; Mehrfachnennung möglich)

Die Auswertung nach positionsspezifischen Gruppen zeigt dabei (s. Abb. 16), dass die Vielfalt der genannten Formaten bei wissenschaftlichen Mitarbeiter und Doktoranden deutlich höher als bei Professoren und Privatdozenten ist. Dieses Ungleichgewicht lässt sich vermutlich durch unterschiedliche Rollen erklären, die die Forschenden in Projekten einnehmen. Während die Professoren und Privatdozenten in der Regel leitende und koordinierende Funktionen übernehmen, sind wissenschaftliche Mitarbeiter und Doktoranden primär für die Be- und Verarbeitung der Forschungsdaten zuständig.

Abb. 16: Wortwolken zu verwendeten Dateiformaten (von links nach rechts: Auswertung für wissenschaftliche Mitarbeiter, Doktoranden, Professoren bzw. Privatdozenten)

Die Anzahl der verwendeten Formate steigt erwartungsgemäß auch im Rahmen der Drittmittelprojekte deutlich gegenüber der Forschung ohne externe Förderung (s. Abb. 17). Dies ist sicherlich erneut auf deren höhere Komplexität zurückzuführen.

Abb. 17: Wortwolken zu verwendeten Dateiformaten (von links nach rechts: Auswertung für Projekte mit und ohne Forschungsförderung)

3.2.4.2 DATENBESCHREIBUNG

Um die Auffindbarkeit und Nachnutzung der Forschungsdaten zu ermöglichen, sollten diese in Bezug auf Inhalt, Kontext und Struktur durch Metadaten (z.B. DataCite oder Dublin Core) beschrieben werden. Dabei ist es essentiell, normierte Begriffe bzw. Vokabulare (z.B. GND oder Iconclass) für eine eindeutige Identifizierung einer Entität (z.B. Person, Ort, Epoche) zu verwenden.

Wie die Ergebnisse der Befragung jedoch zeigen (s. Abb. 18), ist das Bewusstsein für die Notwendigkeit einer standardisierter Beschreibung der Forschungsdaten nicht ausreichend vorhanden. So gaben nur 35% der Befragten an, für die Erschließung mindestens eines Teils ihrer Daten Normdaten und kontrollierte Vokabulare heranzuziehen. Dabei wurden in erster Linie XML-basierte Standards, Getty Thesauri, GND sowie Iconclass genannt. Bemerkenswert ist jedoch, dass lediglich drei Personen angaben, Metadaten-Standards zur Beschreibung ihrer Daten einzusetzen. Wobei die terminologische Unklarheit der Begriffe in diesem Zusammenhang auffallend ist, da bei der Spezifizierung der Antworten in den Freitextkommentaren teilweise widersprüchliche Angaben auftraten.

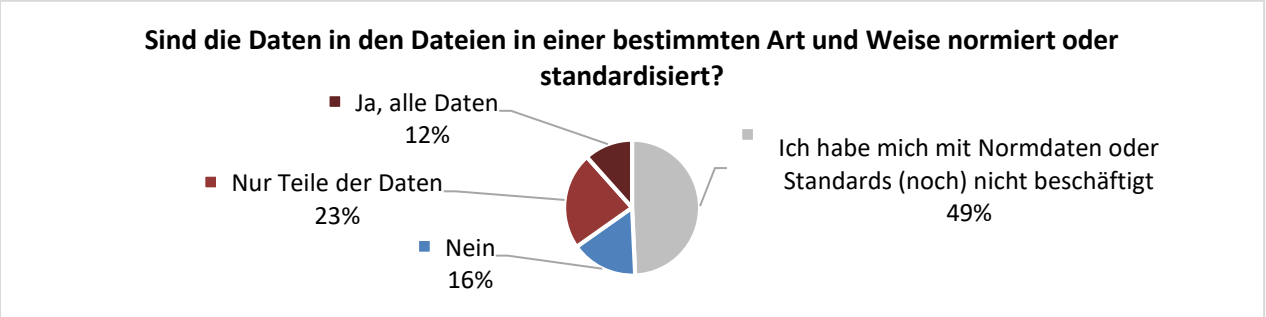

Abb. 18: Datendokumentation (Gesamtauswertung)

Erwartungsgemäß wird auf die Verwendung der Normdaten und Standards im Rahmen von Drittmittelprojekten stärker geachtet. Diese Tendenz ist höchstwahrscheinlich auf die Tatsache zurückzuführen, dass ein solcher Umgang mit Forschungsdaten häufig durch Förderrichtlinien vorgegeben ist.

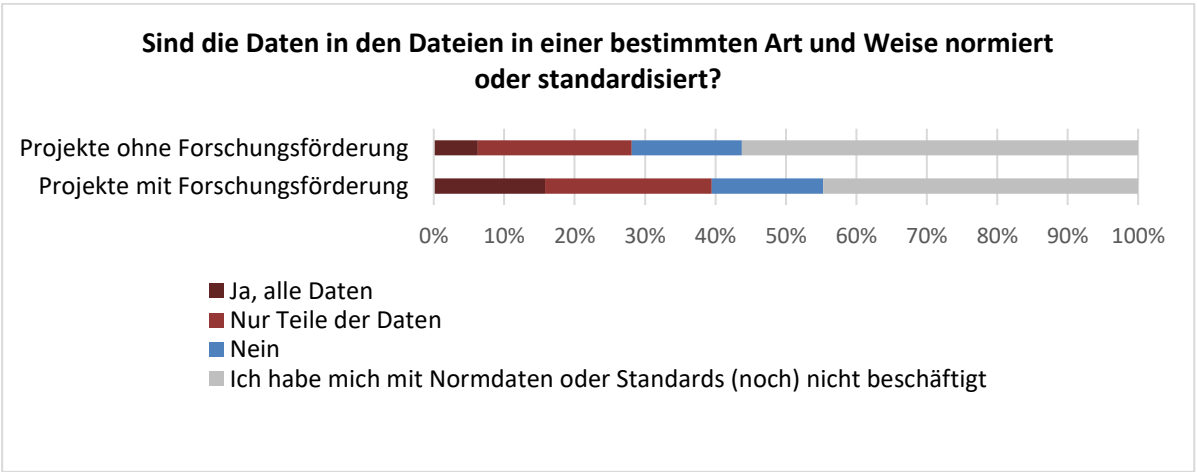

Abb. 19: Datendokumentation (Auswertung für Projekte mit und ohne Forschungsförderung)

3.2.5 Datenspeicherung und -sicherung

Zur Vermeidung des Datenverlustes und der unbefugten Nutzung ist eine sichere und langfristige Speicherung der Forschungsdaten unerlässlich. Bei der Auswahl der Speicherorte sollten daher verschiedene Sicherheitsrisiken beachtet werden. Darüber hinaus ist eine regelmäßige Erstellung von Kopien der Dateien (sog. Backups) gefragt, um sie im Falle eines Hardware- oder Softwaredefekts wiederherstellen zu können.

3.2.5.1 Speicherorte

Die aktuelle Praxis bei der Speicherung der Forschungsdaten zeigt (s. Abb. 20), dass viele Daten lokal auf Rechnern hinterlegt werden. Ferner wird ein erheblicher Teil der Forschungsdaten auf externen Medien wie USB-Stick oder Festplatte sowie auf Servern der Universität gespeichert. Als weitere Speichermedien wurden Cloud-Lösungen genannt, darunter am häufigsten Dropbox, Google Drive und FAUbox. Wie in zahlreichen Freitextkommentaren angegeben wird, sind universitäre Angebote gegenüber den Lösungen kommerzieller Anbieter jedoch meist nicht konkurrenzfähig, weshalb auf letztere verstärkt zurückgegriffen wird.

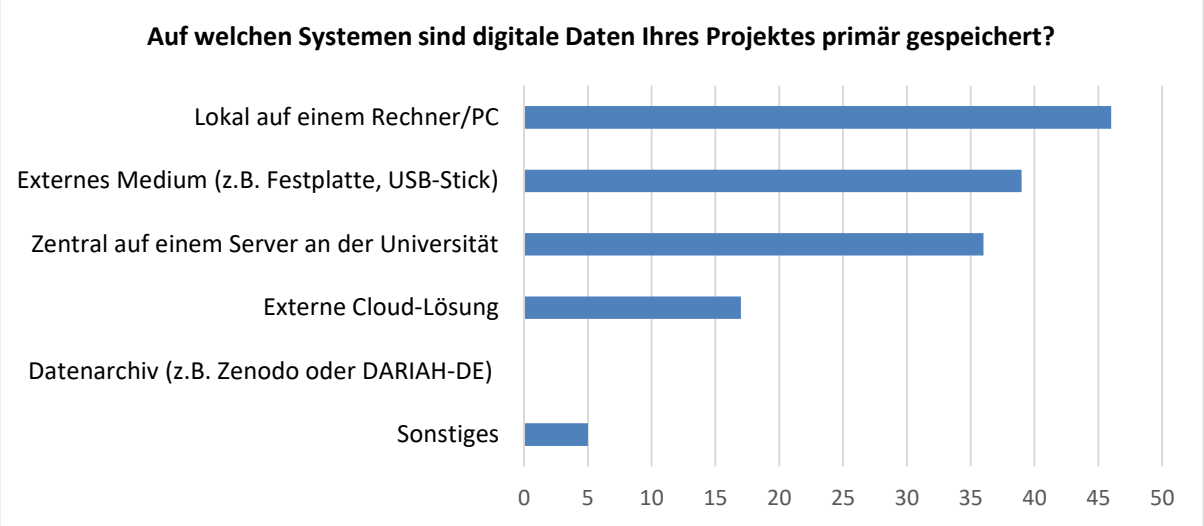

Abb. 20: Speicherorte (Gesamtauswertung; Mehrfachnennungen möglich)

Auffallend ist, dass Cloud-Lösungen eine bedeutend höhere Präferenz bei Doktoranden als bei Professoren und Privatdozenten erfahren (s. Abb. 21). Die Skepsis der letzteren Karrieregruppe gegenüber diesen Speichermedien ist, wie aus den Interviews hervorging, in erster Linie auf die fragwürdige Informationssicherheit der meisten kommerziellen Angebote zurückzuführen, vor allem mit Blick auf den Umgang mit vertraulichen Daten. Als Projektleiter müssen Professoren und Privatdozenten auf Datenschutz und Datensicherheit in besonderem Maße achten, worauf die Sensibilisierung bei der Auswahl der Speichersysteme zurückzuführen ist.

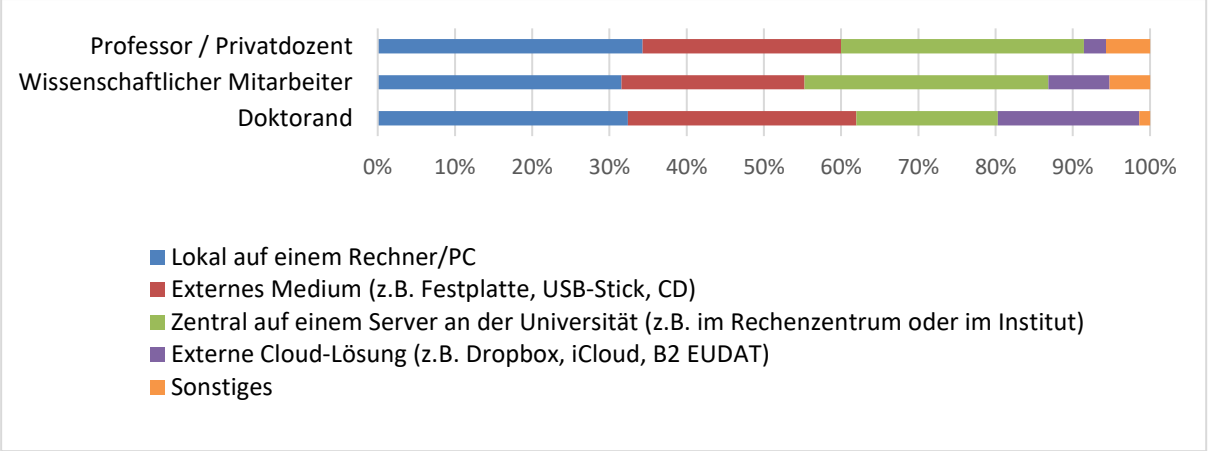

Abb. 21: Speicherorte (Auswertung nach Karrierestufen)

Beachtlich ist zudem, dass keiner der Befragten angab, die Daten in Datenarchiven zu speichern. Dies lässt sich höchstwahrscheinlich dadurch erklären, dass die Möglichkeiten der Datenarchive und Repositorien in den Geistes- und Sozialwissenschaften an der FAU nur wenig bekannt sind, was die Ergebnisse der Online-Umfrage zweifellos bestätigen (s. Abb. 22).

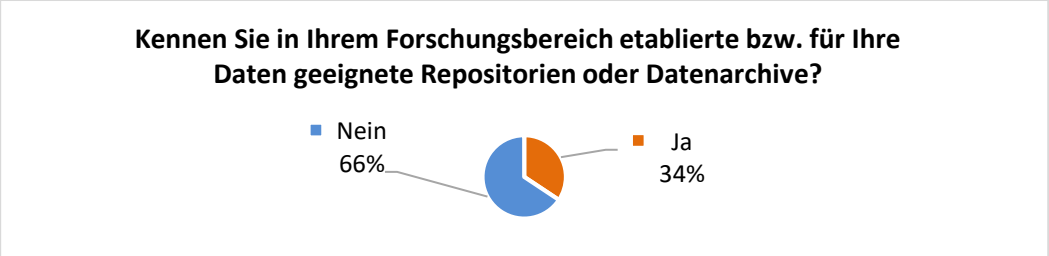

Abb. 22: Wissensstand über Repositorien und Datenarchive (Gesamtauswertung)

3.2.5.2 Backup

Laut der Ergebnisse der Online-Umfrage ist die Bedeutung der Anfertigung von Sicherheitskopien vielen Forschenden bewusst (s. Abb. 23). Knapp über die Hälfte der Befragten (52%) verwies immerhin auf die Erstellung von Backups in sehr kurzen, regelmäßigen Abständen (von täglich bis wöchentlich). Allerdings gab ein erheblicher Teil der Umfrageteilnehmer an, sich unregelmäßig (33%) oder nie (9%) mit der Datensicherung zu beschäftigen. Aufgrund der durchgeführten Interviews kann angenommen werden, dass die Sensibilisierung für diese Thematik insbesondere in datenintensiven Disziplinen bzw. Forschungskontexten auftritt.

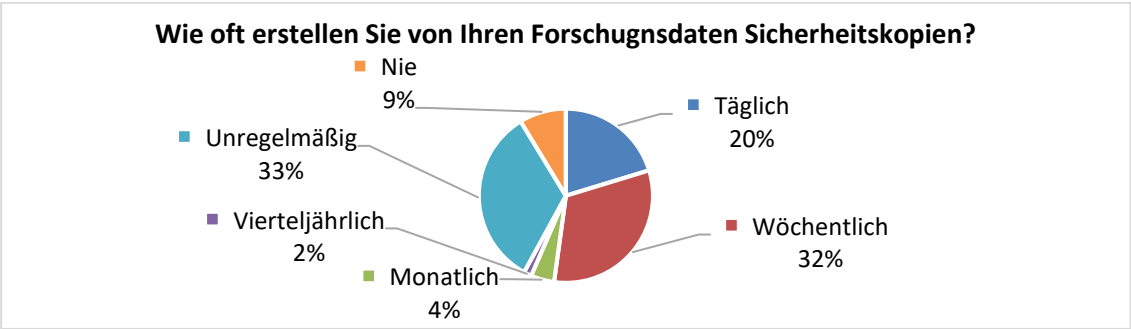

Abb. 23: Erstellung von Sicherheitskopien (Gesamtauswertung)

3.2.6 Datenaufbewahrung

Im Laufe eines Forschungsprozesses entstehen Daten, die strengen Auflagen (z.B. Datenschutz oder Urheberrechte) unterliegen. Hier gilt es, besondere Richtlinien - wie etwa Zugriffsrechte und -restriktionen - bei der Aufbewahrung der Forschungsdaten einzuhalten.

Wie die Ergebnisse der Online-Umfrage vermitteln (s. Abb 24), bestehen in der Tat zahlreiche rechtliche Regelungen, die in den geistes- und sozialwissenschaftlichen Projekten an der FAU beachtet werden müssen. Als solche wurden in den Textkommentaren Datenschutz, Löschungsfristen, Urheberrechte und Lizenzierung am häufigsten genannt. Wie zu erwarten war, berichteten die Teilnehmer der Leitfaden-Gespräche, insbesondere im Umgang mit personenbezogenen Daten Aufbewahrungspflichten nachkommen zu müssen. Auch bei der Durchführung der Drittmittelprojekte treten die Auflagen häufig auf (s. Abb. 25). Diese Entwicklung ist gewiss dadurch zu erklären, dass viele Drittmittelgeber eine gesicherte Aufbewahrung der Daten bei der Vergabe der Förderung explizit fordern.

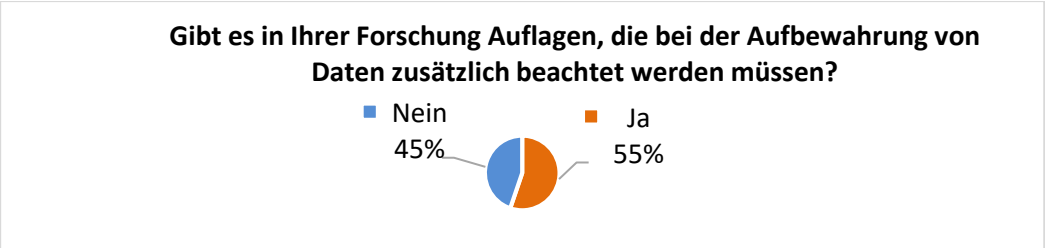

Abb. 24: Auflagen bei der Datenaufbewahrung (Gesamtauswertung)

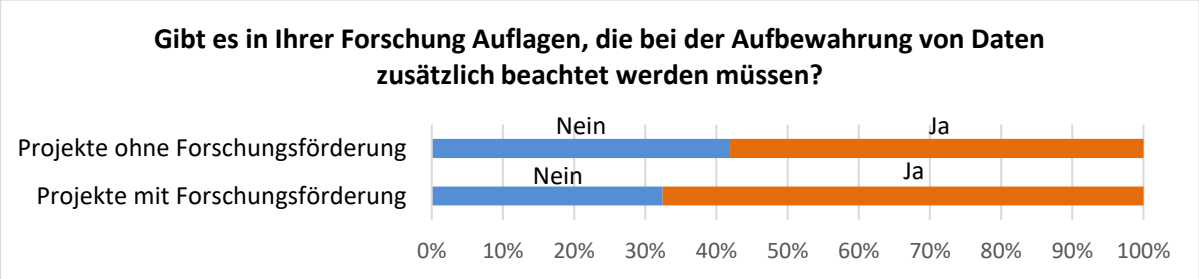

Abb. 25: Auflagen bei der Datenaufbewahrung (Auswertung für Projekte mit und ohne Forschungsförderung)

3.2.7 Zugänglichkeit und Nachnutzung

Die Bereitstellung der Forschungsdaten ist wichtig, da diese im Kontext neuer Fragestellungen nachgenutzt werden können. Außerdem ermöglicht der Zugang zu Daten die Transparenz und die Nachvollziehbarkeit der Forschungsergebnisse.

Die Befunde der Studie belegen jedoch, dass eine Nachnutzung und Veröffentlichung der Forschungsdaten in den meisten Fällen nicht gewährleistet ist. Lediglich 19% der Befragten bejahten die Frage, ob sie ihre Forschungsdaten jemals öffentlich zugänglich gemacht haben (s. Abb. 26).

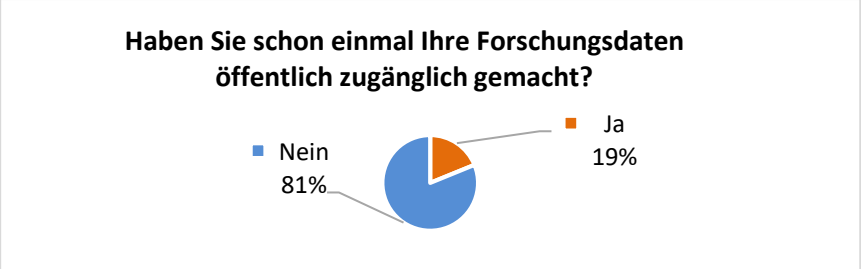

Abb. 26: Zugänglichkeit der Daten (Gesamtauswertung)

Ebenso selten werden Forschungsdaten anderer Wissenschaftler nachgenutzt, wobei das Verhältnis von Nachnutzung der Fremddaten gegenüber der Weitergabe eigener Ergebnisse deutlich positiver ausfällt. So berichteten 37% der Umfrageteilnehmer, mindestens einmal Daten einer anderen Forschungsgruppe nachgenutzt zu haben (s. Abb. 27).

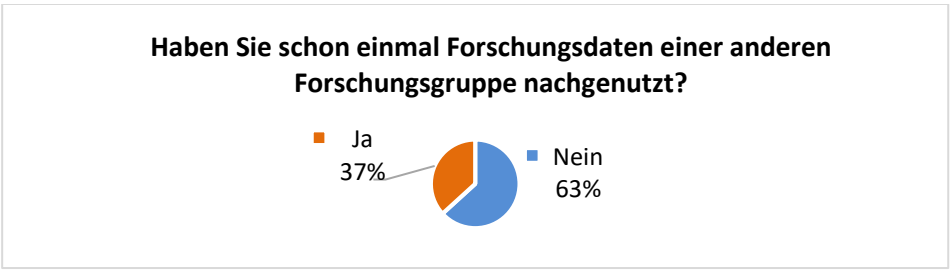

Abb. 27: Nachnutzung der Daten (Gesamtauswertung)

Diese beschriebene Zurückhaltung gegenüber der Bereitstellung und Nachnutzung der Forschungsdaten lässt sich aufgrund der Aussagen der Interviewteilnehmer primär auf rechtliche Schwierigkeiten und Unklarheiten zurückführen, die in den Geistes- und Sozialwissenschaften mehrheitlich vorherrschen. Die Befürchtung von Datenschutzverletzungen hält Forschende häufig davon ab, Forschungsdaten einer breiten Öffentlichkeit zugänglich zu machen bzw. diese nachzunutzen. Auch die fehlende Einwilligung seitens Dritter wurde als Hinderungsgrund hinsichtlich einer Publikation der Forschungsdaten genannt.

Hinzu kommt, dass sich nicht alle Forschungsdaten - wie aus den Interviews ersichtlich wurde - für eine Nachnutzung eignen. Insbesondere betrifft dies die bei einer Datenerhebung unmittelbar gewonnenen, noch unbearbeiteten Roh- bzw. Primärdaten. Aufgrund mangelnder Qualität oder der uneindeutigen Zuordnung zum Objekt, das sie beschreiben, sind sie für andere Akteure kaum interpretierbar. So besitzen beispielsweise originale Luftaufnahmen, die zu einer 3D-Rekonstruktion herangezogen werden, für Sekundärforscher keine wissenschaftliche Relevanz im Sinne einer Informations- und Interpretationsquelle. Ebenfalls kritisch wird eine sekundäranalytische Auswertung der qualitativen Datensätze gesehen, die im Zuge kontextgebundener Forschungsprozesse entstehen. Hierzu zählen unter anderem für Zwecke empirischer Sozialforschung erhobene Interviews und Umfragen, da sie oft mit einer spezifischen Fragestellung verknüpft und somit im Kontext eines anderen Forschungsschwerpunktes kaum auswertbar sind.

3.3 Anforderungs- und Bedarfsanalyse

Im Folgenden werden Anforderungen und Bedürfnisse der Forscher aus den Geistes- und Sozialwissenschaften an der FAU mit Blick auf den Umgang mit digitalen Forschungsdaten näher beleuchtet.

3.3.1 Anforderungen an ein Datenmanagementplan-Tool

Neben anderen Maßnahmen zur Unterstützung der Wissenschaftler im Umgang mit ihren Forschungsdaten wird im Rahmen des Projektes “eHumanities-interdisziplinär” ein DMP-Werkzeug(konzept) entwickelt. Dieses Tool soll anschließend bei der Planung der strukturierten Datenaufbereitung und -dokumentation in den Geistes- und Sozialwissenschaften eingesetzt werden. Um dabei den Bedürfnissen der Zielgruppe gerecht zu werden, wurde im Rahmen der Bedarfserhebung die Frage nach den Anforderungen an ein DMP-Hilfstool gestellt (s. Abb. 28).

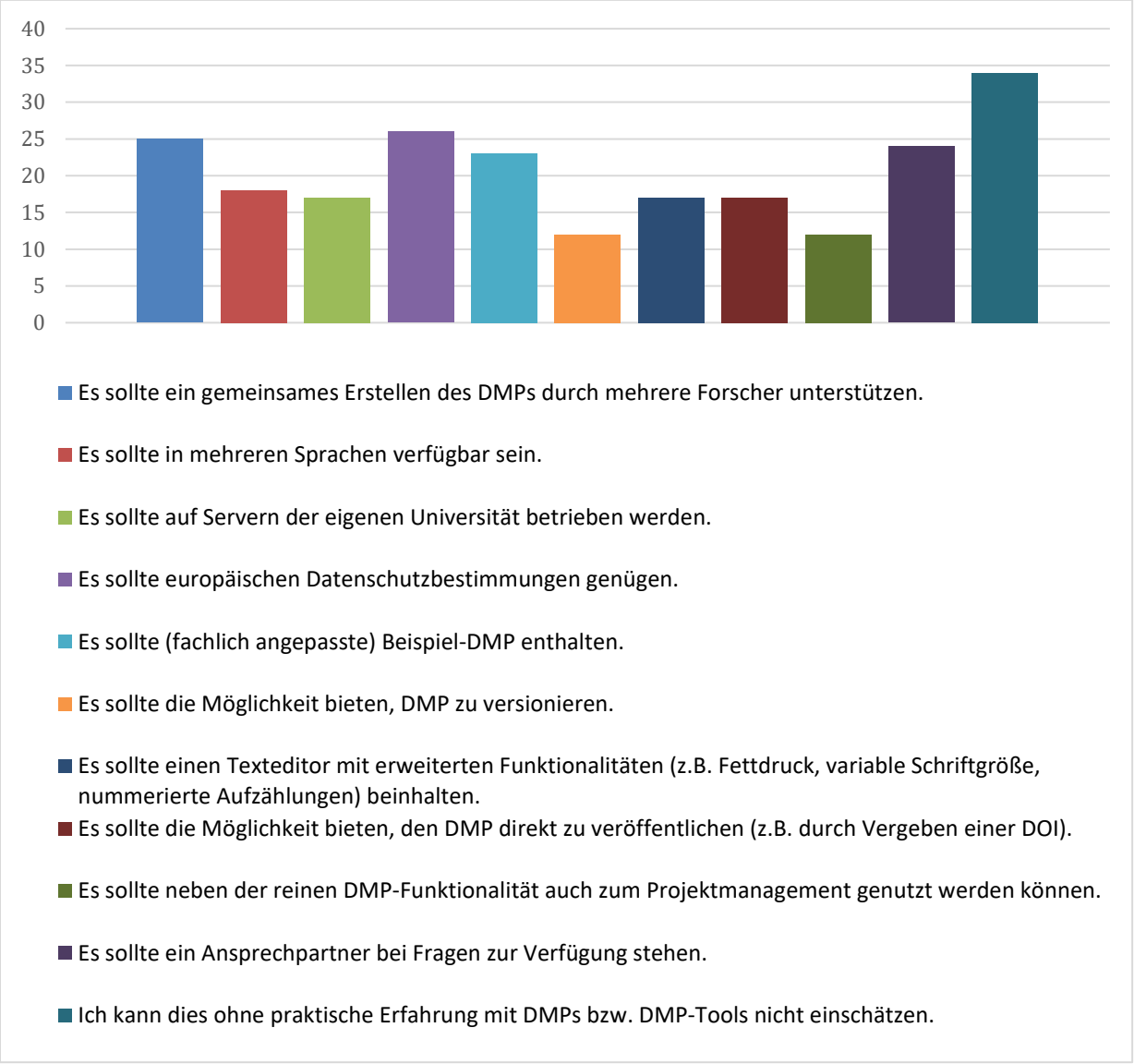

Abb. 28: Anforderungen an ein DMP-Hilfstool (Gesamtauswertung; Mehrfachnennungen möglich)

Es stellte sich heraus, dass die Mehrheit der Befragten ihre Wünsche hinsichtlich eines DMP-Tools aufgrund der fehlenden Praxis im Umgang mit diesen nicht äußern konnte. Die Antworten anderer Umfrageteilnehmer führten zu der Schlussfolgerung, dass die Möglichkeiten zum Erstellen eines DMPs durch mehrere Forscher, die Vereinbarkeit mit den europäischen Datenschutzbestimmungen, das Vorhandensein einzelner Fallbeispiele und von Ansprechpartnern zu den wichtigsten Anforderungen der Zielgruppe gehören.

Von den weiteren vorgegebenen Auswahloptionen wurden dagegen die Möglichkeiten zum Einsatz für das Projektmanagement und zur Versionsverwaltung am seltensten genannt. Dies lässt sich womöglich dadurch erklären, dass gezielte Managementmethoden - den Interviews zufolge - in den Geisteswissenschaften nur selten eingesetzt werden. Überraschend ist, dass die Möglichkeit zur Versionierung nur für wenige Forscher von Bedeutung ist. Wie aus den Leitfadengesprächen jedoch hervorging, ist die Dynamik bei einer Planerstellung in einem Forschungsvorhaben essentiell. Der Grund dafür ist, dass Forschung immer an den Grenzen des momentanen Wissens agiert und unvorhersehbare Entwicklungen daher natürlicher Bestandteil des Forschungsprozesses sind. Infolgedessen besteht häufig ein Änderungsbedarf am ursprünglichen Plan. Dieser Bedarf kann allerdings mit der Funktionalität der Versionierung abgedeckt werden. Aufgrund der mangelnden praktischen Erfahrungen mit DMP-Tools ist diese Tatsache jedoch den meisten Wissenschaftlern offensichtlich nicht bewusst.

3.3.2 Bedarf nach Veröffentlichung bzw. Zugänglichmachung

Zur Ermittlung der Bewusstseinsbildung für Möglichkeiten der Nachnutzung von Forschungsdaten wurden im Rahmen der Studie Fragen nach dem Bedarf an Veröffentlichung und Zugänglichmachung der Daten gestellt.

Die Ergebnisse der Online-Umfrage zeigen (s. Abb. 29), dass die grundsätzliche Bereitschaft, eigene Forschungsdaten anderen zur Verfügung zu stellen, in den Geistes- und Sozialwissenschaften an der FAU weitgehend vorhanden ist. Lediglich sechs Teilnehmer sprachen sich gegen die Weitergabe ihrer wissenschaftlichen Daten aus. Die Befürwortung von Open Access wurde auch im Laufe von Interviews mehrfach bestätigt.

Wie allerdings aus der Online-Umfrage hervorging, setzt die Mehrheit der Forschenden hierfür voraus, dass das jeweilige Projekt zum Zeitpunkt der Datenveröffentlichung bereits abgeschlossen ist. Dies lässt sich aufgrund der Bedenken vor Fehlinterpretationen oder Missbrauch der Forschungsdaten begründen. Eine vergleichsweise wichtige Rolle spielen bei der Entscheidung für oder gegen eine mögliche Datenpublikation auch die Kostenfrage und der technisch-administrative Aufwand. Als weitere wichtige Faktoren wurden in diesem Zusammenhang die Möglichkeit zur Aktualisierung der Daten sowie die Verfügbarkeit eines in der Fachcommunity etablierten Archivs oder Repositoriums erwähnt. Die häufige Nennung der letzteren Voraussetzung bestätigt erneut, dass Geistes- und Sozialwissenschaftler an der FAU nur begrenzt über Informationen hinsichtlich der Datenarchive und Repositorien verfügen, was wiederum die seltene Verwendung dieser wichtigen Datenablageorte erklärt.

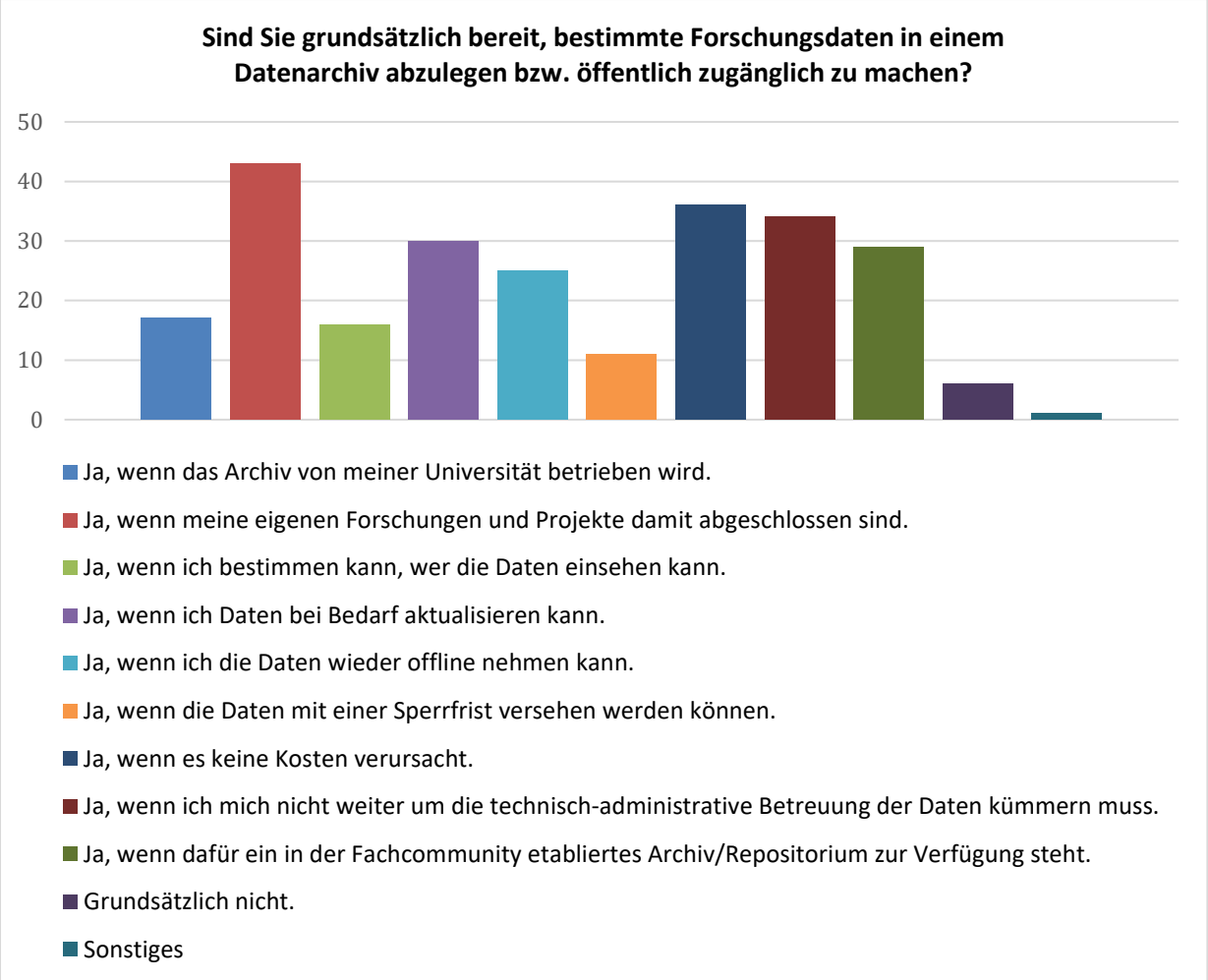

Abb. 29: Bereitschaft zur Ablage der Forschungsdaten in einem Datenarchiv (Gesamtauswertung; Mehrfachnennung möglich)

Werden Forschungsdaten veröffentlicht, so gab die Mehrheit der Umfrageteilnehmer an, diese über Bibliothekskataloge bereitstellen zu wollen (s. Abb. 30). Suchmaschinen und (Fach-)Datenbanken spielen in diesem Zusammenhang ebenfalls eine wichtige Rolle. Hingegen wurden Discovery Services, Verlagsplattformen sowie Fachinformationsdienste deutlich seltener erwähnt. Bei der fachspezifischen Auswertung ist es jedoch beachtlich, dass Fachinformationsdienste trotz eines scheinbar geringen Gesamtstellenwerts eine wichtige Rolle in den Forschungsbereichen der Kunstgeschichte und Medienwissenschaften spielen (s. Abb. 31). Dies ist vermutlich der besonders großen Bedeutung von *arthistoricum.net* innerhalb der kunsthistorischen Forschung zuzurechnen, was anhand der Spezifizierung der Angaben in den Freitextkommentaren festzustellen ist. Andere Fachinformationsdienste scheinen dagegen in den Geistes- und Sozialwissenschaften an der FAU keine vergleichbare Relevanz bzw. Akzeptanz zu besitzen.

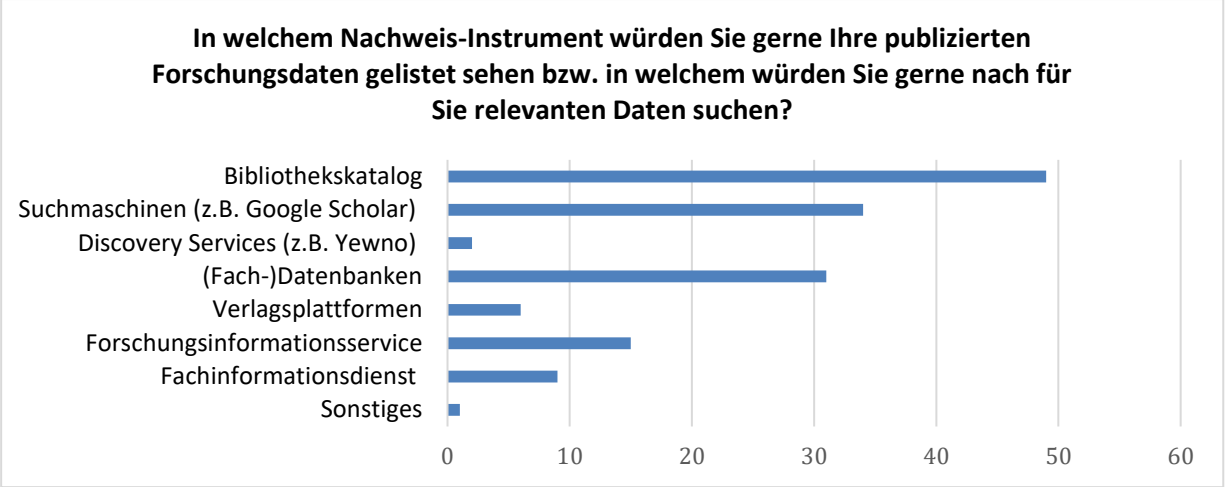

Abb. 30: Publikation der Forschungsdaten (Gesamtauswertung; Mehrfachnennungen möglich)

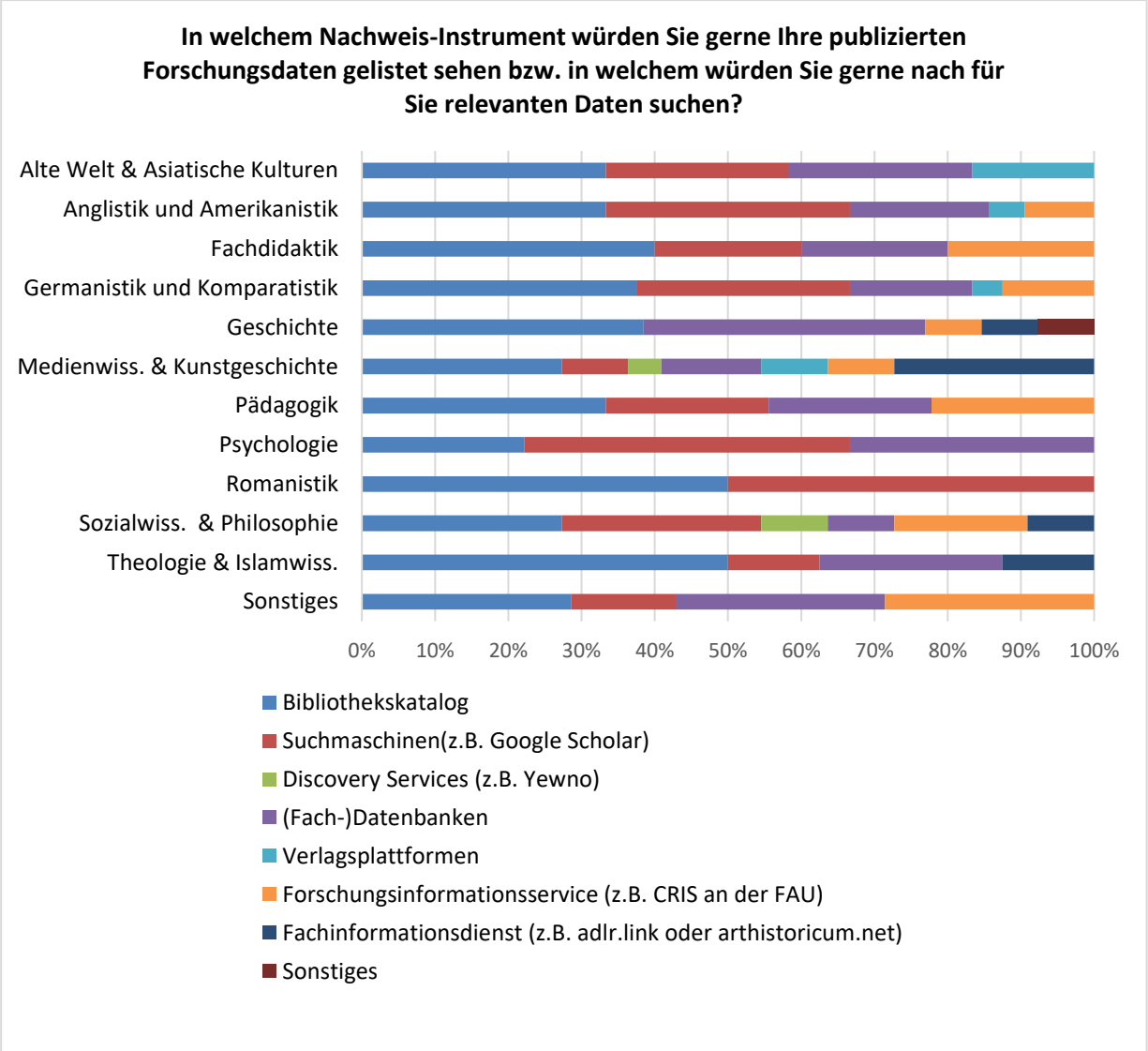

Abb. 31: Publikation der Forschungsdaten (Auswertung nach Fächergruppen)

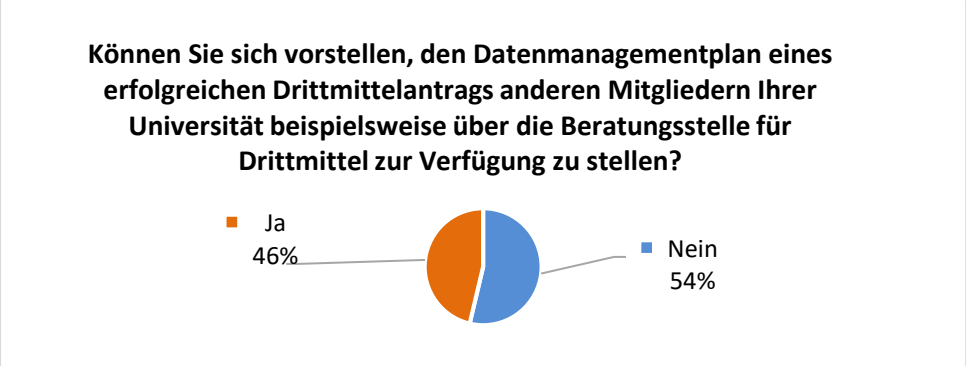

Abb. 32: Bereitschaft zur Weitergabe von DMP (Gesamtauswertung)

Während die Bereitschaft zur Bereitstellung eigener Forschungsdaten verstärkt vorhanden ist, wird die Weitergabe eines erfolgreichen Datenmanagemetplans deutlich kritischer betrachtet, auch wenn die Tendenz mehrheitlich positiv ausfällt (s. Abb. 32). Dabei ist auffallend, dass mit dem Aufstieg in der akademischen Laufbahn die Skepsis gegenüber der Verwendung des eigenen DMPs durch Dritte deutlich nachlässt (s. Abb. 33). Während sich beinahe 40% der Doktoranden in diesem Zusammenhang negativ äußerten, waren es bei wissenschaftlichen Mitarbeitern 25% und bei Professoren gar nur noch 5%. Die Unterschiede bei dieser positionsspezifischen Auswertung sind vermutlich auf den verstärkten Wettbewerb und Konkurrenzdruck zu Beginn einer wissenschaftlichen Karriere zurückzuführen.

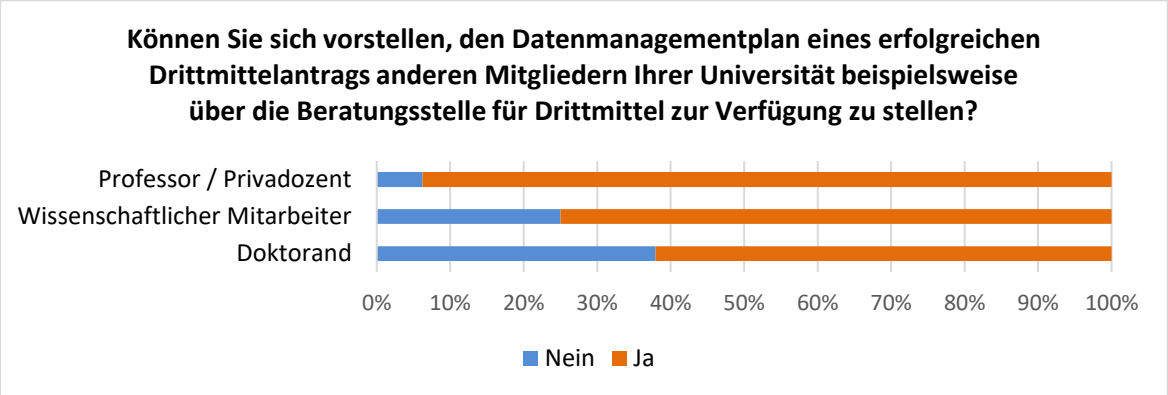

Abb. 33: Bereitschaft zur Weitergabe von DMP (Auswertung nach Karrierestufen)

3.3.3 Beratungs- und Unterstützungsbedarf

Um konkrete Handlungsfelder in Hinblick auf weitere Schulungs- und Informationsangebote möglichst zielgruppenorientiert festzulegen, wurde im Rahmen der durchgeführten Studie der Frage nach Beratungs- und Unterstützungsbedarf nachgegangen. Folgende Themen ließen sich dabei aus Sicht der Forschenden als besonders wichtig festhalten:

Technische Aspekte von FDM wurden als eine der größten Herausforderungen für den Forschungsalltag genannt. Dies betrifft unter anderem den Umgang mit Dateiformaten, Datenbanken, Software oder Tools. In diesem Zusammenhang wiesen viele Interviewteilnehmer auf Zeit-

mangel im wissenschaftlichen Arbeiten hin, infolgedessen die Suche nach effizienteren Herangehensweisen nur in Einzelfällen vorgenommen wird. Zur Vermeidung eines zusätzlichen Zeitaufwandes wird somit in der Regel auf pragmatische, jedoch mehrheitlich kurzfristige Lösungsansätze zurückgegriffen.

Eine große Bedeutung wurde auch rechtlichen Aspekten bei digitalen Forschungsdaten zugeschrieben. Die Forschenden wiesen dabei auf die Unklarheit hin, die häufig im Umgang mit rechtlichen Themen - insbesondere hinsichtlich der Fragen nach Lizenzen und Urheberrechten - herrscht. Diese Unsicherheiten haben oft zur Folge, dass viele Forschungsprojekte nicht in der gewünschten Form realisiert werden können. Um Copyright-Probleme gar nicht erst entstehen zu lassen, werden zahlreiche technische Möglichkeiten nicht ausgenutzt.

Die Nutzung von Metadaten und Standards wurde ebenfalls als ein wichtiges FDM-Thema genannt. Der Umgang mit diesem Sachverhalt scheint größtenteils auf der Ebene der elementaren Kompetenzen zu erfolgen. Ein tiefergehendes Verständnis für Metadaten ist jedoch nicht selten gefragt. Insbesondere betrifft dies Forschungskontexte, in denen die Arbeit mit digitalen Korpora und Datenbanken im Vordergrund steht - eine Entwicklung, die bei den an der Philosophischen Fakultät der FAU vertretenen Disziplinen deutlich an Fahrt aufnimmt.

Um den genannten Herausforderungen entgegenzutreten, erachtet es die Mehrheit der Interviewteilnehmer für wichtig, mit der entsprechenden Kompetenzvermittlung bereits zwischen dem Abschluss eines grundständigen Studiums und dem Anfang eines Promotionsvorhabens zu beginnen. Allerdings sehen die Wissenschaftler die Rolle des Informations- und Wissensvermittlers (Berater, Coach) größtenteils bei den universitätsinternen Forschungsdienstleistern wie der Universitätsbibliothek oder dem Rechenzentrum. Der Hauptgrund hierfür ist die Skepsis hinsichtlich der möglichen Einbindung der FDM-Themen in die bestehenden Curricula, da reguläre Lehrveranstaltungsformate mit anderen Inhalten und Verpflichtungen bereits stark angefüllt sind. Zwar werden fachspezifische Aspekte im Rahmen einiger Studiengänge thematisiert, setzen jedoch entsprechende Kernkompetenzen voraus, die aus Zeitgründen im Studium nicht angesprochen werden.

4. Fazit

Im Folgenden werden die wichtigsten Erkenntnisse zusammengefasst bzw. weitgehend verbreitete Defizite und Unsicherheiten in Kürze erläutert. Diese gilt es im Rahmen der Kompetenzvermittlung und der Durchführung von Qualitätssicherungsmaßnahmen mit Blick auf effizientes FDM in den Geistes- und Sozialwissenschaften der FAU zukünftig aufzuarbeiten.

Nicht zuletzt aufgrund der Fächer- und Disziplinvielfalt an der Philosophischen Fakultät ist es jedoch ausgesprochen schwierig, allgemeingültige Aussagen zu treffen. Die Befragung der Forschenden aus allen vertretenen Teildisziplinen war im Rahmen der durchgeführten Studie nicht möglich. Ein weiteres Hindernis hinsichtlich allgemeingültiger Rückschlüsse stellt die geringe

Beteiligung der angesprochenen Wissenschaftler an der Online-Umfrage dar. Vor dem Hintergrund der beschriebenen Ausgangssituation muss die Aussagekraft der Evaluationsergebnisse relativiert werden. Infolgedessen gilt es, die Befunde der vorliegenden Studie vielmehr als Hypothesen aufzufassen.

Zunächst kann positiv festgehalten werden, dass das Bewusstsein für den Mehrwert einer nachhaltigen Verfügbarkeit und interdisziplinären Nachnutzbarkeit von Forschungsdaten in den Geistes- und Sozialwissenschaften an der FAU stark ausgeprägt ist. Die grundsätzliche Bereitschaft, eigene Forschungsdaten anderen zur Verfügung zu stellen, ist weitgehend vorhanden.

Allerdings ist anzumerken, dass der Begriff Forschungsdaten oft ausschließlich mit statistischen Daten oder mit den Naturwissenschaften in Verbindung gebracht wird. In den Geistes- und Sozialwissenschaften werden anstelle "Forschungsdaten" häufig Begriffe wie Primärdaten, Grundlagendaten, Quellen oder Material verwendet. Aus diesem Grund fühlten sich vermutlich viele Forschende durch die Online-Umfrage nicht angesprochen, da sie ihre wissenschaftliche Tätigkeit nicht mit dem Begriff Forschungsdaten assoziieren. Ein großer Bedarf an Sensibilisierung besteht aber vor allem in Fragen der konkreten Handhabung der Daten über ihren gesamten „Lebenszyklus" hinweg. Größtenteils treten dabei folgende fachübergreifende Problemstellungen auf:

- **Planung der Datenerhebung:** In der Vorbereitungsphase eines Forschungsvorhabens erfolgt eine strukturierte Planung des FDMs nur in Einzelfällen. In der Regel werden lediglich allgemeine Überlegungen getroffen, aufgrund derer das frühzeitige Erkennen von Fehlern und Schwachstellen nur begrenzt möglich ist.

- **Aufbereitung der Daten:** Bei der Aufbereitung der anfallenden Forschungsdaten wird nicht durchgehend auf die Qualität und Möglichkeiten gewählter Dateiformate geachtet. Die Nutzung von proprietären Dateiformaten ist immer noch weit verbreitet. Auch Formate der strukturierten Datenaufbereitung werden aktuell nur bedingt genutzt. Diese Entwicklung ist primär auf fehlende Kompetenzen im Umgang mit Dateiformaten und Software zurückzuführen. Die meisten Forschenden sind nur begrenzt über Alternativmöglichkeiten hinsichtlich der üblicherweise verwendeten Dateiformate und Software informiert. Auch im Umgang mit bekannten Anwendungen fehlt häufig die Expertise, um technologische Leistungsfähigkeit und Grenzen der eingesetzten Formate (z.B. Vor- und Nachteile von JPEG, PNG und TIFF bzw. CSV und XLS) richtig einzuschätzen.

- **Beschreibung der Daten:** Metadaten werden zur Datenbeschreibung gegenwärtig lediglich in Ausnahmefällen eingesetzt. Nicht konsequent wird auf die Verwendung standardisierter Vokabulare geachtet. Die terminologische Unklarheit der Begriffe Metadaten und Normdaten ist dabei auffallend.

- **Dokumentation der Daten:** Die Entstehung eines Datensatzes wird ebenfalls nur in Einzelfällen dokumentiert. Dies führt häufig dazu, dass unerwartet aufgetretene Ungereimtheiten

nicht nachvollzogen werden können, da etwa der ursprüngliche Hersteller der Daten nicht mehr im Projekt tätig ist. Insbesondere bei langfristigen Projekten kann dies der Fall sein.

- **Archivierung der Daten:** Ein erheblicher Teil der Forschungsdaten wird weiterhin auf lokalen Speichersystemen in unorganisierter Form aufbewahrt, sodass sie nach einer gewissen Zeit in Vergessenheit geraten oder sogar ganz gelöscht werden. Ferner werden für die Datenablage, wenn auch in geringerem Umfang, Cloud-Lösungen kommerzieller Anbieter verwendet, die dem nationalen bzw. europäischen Datenschutzrecht nicht unterliegen.

- **Bereitstellung der Daten:** Trotz des meist vorhandenen Bewusstseins um die Bedeutung der Datennachnutzung, ist die Mehrheit der Forschenden weitgehend zurückhaltend, was die Weitergabe eigener wissenschaftlicher Daten betrifft. Dies ist primär auf Unsicherheiten im Umgang mit rechtlichen Fragen zurückzuführen. Fehlende Kompetenzen beziehen sich dabei in erster Linie auf Unklarheiten in Bezug auf konkrete Anforderungen (beispielsweise ob ein bestimmtes Bild in einem E-Book publiziert werden darf), wobei eine deutlich stärkere Sensibilisierung für rechtliche Risiken bei Studierenden und Nachwuchswissenschaftlern wünschenswert ist. Ein weiteres Hindernis mit Blick auf die Weitergabe eigener Forschungsdaten stellt die fehlende Kompetenz im Umgang mit Repositorien und Datenarchiven dar, die an der Philosophischen Fakultät der FAU nicht vollumfänglich bekannt sind. Selbst wenn Forschende über entsprechende Informationen verfügen, können sie mit diesen Infrastrukturen nur begrenzt umgehen.

Es ist anzunehmen, dass die mangelnde Sensibilisierung im Umgang mit wissenschaftlichen Daten insbesondere in den Forschungskontexten besteht, in denen Erkenntnisse aus überlieferten Quellen bzw. aus einer überschaubaren Menge an Forschungsdaten gewonnen werden. Dies betrifft in der Regel Disziplinen, in denen hermeneutisch-interpretierende Verfahren im Vordergrund stehen. Dagegen müssen auf Empirie gestützte Wissenschaftler ihre quantitativen und qualitativen Daten größtenteils grundsätzlich neu erheben, womit eine gewisse Sensibilisierung etwa bei der Auswahl der Dateiformate und Software einhergeht. Auch in Disziplinen, in denen kontinuierlich hohe Datenmengen anfallen, wird die Bewusstseinsbildung für FDM-Themen stärker gefördert, da allein die Verwaltung und Archivierung der Daten anspruchsvolle Aufgaben darstellen.

Nicht zu unterschätzen ist zudem die Rolle der Förderinitiativen bei der Schaffung eines Bewusstseins der Forschenden für das Thema. Durch entsprechende Regularien und Richtlinien fordern sie explizit bei der Vergabe der Mittel einen verantwortungsvollen Umgang mit wissenschaftlichen Daten. Dieser positive Einfluss der Drittmittelgeber zeigte sich deutlich in den Ergebnissen der vorliegenden Studie.