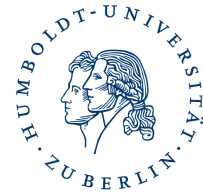


HUMBOLDT-UNIVERSITÄT ZU BERLIN



Forschungsdaten an der Humboldt-Universität zu Berlin

Bericht über die Ergebnisse der Umfrage zum Umgang mit digitalen Forschungsdaten an der Humboldt-Universität zu Berlin

Version 1.0

Elena Simukovic, Maxi Kindling, Prof. Dr. Peter Schirmbacher

Oktober 2013

Zitationsvorschlag:

Simukovic, Elena¹; Kindling, Maxi²; Schirnbacher, Peter³ (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](https://nbn-resolving.org/urn:nbn:de:kobv:11-100213001)

Computer- und Medienservice
Institut für Bibliotheks- und Informationswissenschaft
Humboldt-Universität zu Berlin

Schlagwörter:

Forschungsdaten, Forschungsdatenmanagement, Umfrage, Gute wissenschaftliche Praxis

Dieser Bericht ist auf Grundlage einer Umfrage an der Humboldt-Universität zu Berlin entstanden. Eine anonymisierte Auswertung der Ergebnisse und erläuternde Tabellendokumente stehen zusätzlich zur Verfügung (s. Literaturverzeichnis).

Zur besseren Lesbarkeit wird im Bericht ausschließlich die männliche Form verwendet. Es sollen sich jedoch beide Geschlechter gleichermaßen angesprochen fühlen.



Dieses Werk bzw. Inhalt steht unter einer [Creative Commons Namensnennung 3.0 Deutschland Lizenz](https://creativecommons.org/licenses/by/3.0/de/).

1 ORCID: [0000-0003-1363-243X](https://orcid.org/0000-0003-1363-243X)
2 ORCID: [0000-0002-0167-0466](https://orcid.org/0000-0002-0167-0466)
3 ORCID: [0000-0003-2418-6121](https://orcid.org/0000-0003-2418-6121)

Inhaltsverzeichnis

Glossar.....	6
0 Vorwort.....	7
1 Einleitung	7
1.0 Motivation und Hintergrund.....	7
1.1 Forschungsdatenmanagement an Hochschulen.....	8
1.2 Ausgangslage der Humboldt-Universität zu Berlin.....	9
1.2 Forschungsdatenmanagement-Initiative.....	10
2 Methodik.....	10
2.1 Gestaltung des Fragebogens.....	10
2.2 Durchführung der Umfrage.....	11
3 Ergebnisse.....	12
3.1 Teilnehmer und Rücklauf.....	12
3.1.1 Einrichtungen.....	12
3.1.2 Teilnehmergruppen.....	13
3.2 Datentypen.....	14
3.2.1 Quellen.....	14
3.2.2 Typen von Forschungsdaten.....	14
3.2.3 Fach- oder gerätespezifische Daten.....	15
3.2.4 Speicherorte.....	16
3.2.5 Sicherung auf Datenträgern.....	16
3.2.6 Gesamtgröße der Forschungsdaten.....	17
3.2.7 Erstellung der Sicherungskopien.....	18
3.2.8 Verantwortung für Forschungsdaten.....	18
3.3 Zugang.....	19
3.3.1 Zugriff auf Forschungsdaten.....	19
3.3.2 Veröffentlichung und Nachnutzung.....	19
3.3.3 Sicherung guter wissenschaftlicher Praxis.....	21
3.3.4 Weitere Richtlinien.....	23
3.4 Infrastruktur und Service.....	23
3.4.1 Bereitschaft zur Zugänglichmachung.....	23
3.4.2 Forschungsdaten in einem Datenarchiv.....	24
3.4.3 Bevorzugtes Datenarchiv.....	25
3.4.4 Hilfreiche Kriterien.....	27
3.4.5 Nutzung anderweitiger Infrastruktur.....	30
3.4.6 Gewünschte Unterstützung.....	30
3.5 Weiteres Interesse.....	32
4 Vergleich der Antworten zwischen Teilnehmergruppen.....	33
4.1 Datentypen.....	33
4.2 Zugang.....	35
4.3 Infrastruktur und Service	38
4.4 Weiteres Interesse.....	43
5 Zusammenfassung.....	44
6 Literaturverzeichnis.....	47
7 Anhänge.....	50
7.1 Einladungstext (Deutsche Version).....	50
7.2 Fragebogen (Deutsche Version).....	51

Abbildungsverzeichnis

Abbildung 1: Verteilung der Teilnehmer nach Einrichtungen	12
Abbildung 2: Teilnehmergruppen.....	13
Abbildung 3: Quellen von Forschungsdaten.....	14
Abbildung 4: Typen von Forschungsdaten.....	15
Abbildung 5: Fach- oder gerätespezifische Daten.....	15
Abbildung 6: Speicherorte.....	16
Abbildung 7: Sicherung auf Datenträgern.....	17
Abbildung 8: Gesamtgröße der Forschungsdaten.....	17
Abbildung 9: Erstellung der Sicherungskopien.....	18
Abbildung 10: Verantwortung für Forschungsdaten.....	18
Abbildung 11: Zugriff auf Forschungsdaten.....	19
Abbildung 12a: Veröffentlichung und Nachnutzung: In einem Datenarchiv abgelegt?.....	20
Abbildung 12b: Veröffentlichung und Nachnutzung: Heruntergeladen oder zitiert?.....	20
Abbildung 12c: Veröffentlichung und Nachnutzung: Eingereicht oder nachgewiesen?.....	20
Abbildung 13: Veröffentlichung und Nachnutzung (Matrix).....	21
Abbildung 14: Sicherung guter wissenschaftlicher Praxis.....	22
Abbildung 15: Weitere Richtlinien zum Umgang mit Forschungsdaten.....	23
Abbildung 16: Bereitschaft zur Zugänglichmachung.....	24
Abbildung 17: Forschungsdaten für Ablage in einem Datenarchiv.....	25
Abbildung 18: Bevorzugtes Datenarchiv.....	26
Abbildung 19a: Hilfreiche Kriterien: Empfehlung der Kollegen.....	27
Abbildung 19b: Hilfreiche Kriterien: Empfehlung der Fachorganisation.....	27
Abbildung 19c: Hilfreiche Kriterien: Auswertung oder Ranking.....	28
Abbildung 19d: Hilfreiche Kriterien: Bekanntheit der Organisation.....	28
Abbildung 19e: Hilfreiche Kriterien: Entscheide selbstständig.....	28
Abbildung 20: Hilfe bei Entscheidungsfindung (Matrix).....	29
Abbildung 21: Nutzung anderweitiger Infrastruktur.....	30
Abbildung 22: Gewünschte Unterstützung.....	31
Abbildung 23: Gesamtgröße der Forschungsdaten (Vergleich).....	33
Abbildung 24: Erstellung der Sicherungskopien (Vergleich).....	34
Abbildung 25: Verantwortung für Forschungsdaten (Vergleich).....	34
Abbildung 26: Zugriff auf Forschungsdaten (Vergleich).....	35
Abbildung 27a: In einem Datenarchiv abgelegt? (Vergleich).....	36
Abbildung 27b: Heruntergeladen oder zitiert? (Vergleich).....	36
Abbildung 27c: Eingereicht oder nachgewiesen? (Vergleich).....	37
Abbildung 28: Sicherung guter wissenschaftlicher Praxis (Vergleich).....	37
Abbildung 29: Weitere Richtlinien zum Umgang mit Forschungsdaten (Vergleich).....	38
Abbildung 30: Bereitschaft zur Zugänglichmachung (Vergleich).....	39
Abbildung 31: Forschungsdaten für Ablage in einem Datenarchiv (Vergleich).....	39
Abbildung 32: Bevorzugtes Datenarchiv (Vergleich).....	40
Abbildung 33a: Empfehlung der Kollegen (Vergleich).....	40
Abbildung 33b: Empfehlung der Fachorganisation (Vergleich).....	41
Abbildung 33c: Auswertung oder Ranking (Vergleich).....	41
Abbildung 33d: Bekanntheit der Organisation (Vergleich).....	42
Abbildung 33e: Entscheide selbstständig (Vergleich).....	42
Abbildung 34: Gewünschte Unterstützung (Vergleich).....	43

Tabellenverzeichnis

Tabelle 1: Abgleich der Teilnehmergruppen mit der HU-Personalstatistik.....	13
Tabelle 2: Teilnehmerzahl nach Gruppen.....	33

Glossar⁴

Forschungsdaten (eng. *Research data*) – alle Daten, die im Laufe einer wissenschaftlichen Tätigkeit entstehen. Diese können sowohl qualitativer als auch quantitativer Art / Natur sein, sowie in einer digitalen und analogen Form vorliegen. Forschungsdaten stellen in der Regel die Grundlage einer wissenschaftlichen Publikation dar, können aber auch eigenständig publiziert werden.

Forschungsdatenmanagement (eng. *Research data management*) – alle Aktivitäten, die mit der Aufbereitung, Speicherung, Archivierung und Veröffentlichung von Forschungsdaten verbunden sind.

Forschungsdatenmanagementplan (eng. *Data management plan*) – eine kurze Beschreibung der im Laufe eines bestimmten Vorhabens entstehenden Forschungsdaten und der damit verbundenen Forschungsdatenmanagement-Aktivitäten. Ein Forschungsdatenmanagementplan wird oft von Förderorganisationen bei der Beantragung der Fördermittel verlangt. Er kann dabei helfen, bereits in der Planungsphase eines Forschungsvorhabens grundlegende Fragen und Zuständigkeiten zu klären, und ist daher für jede Art der Forschungstätigkeit zu empfehlen.

4 Dieses Glossar liefert die Arbeitsdefinitionen für den vorliegenden Bericht. Sie erfolgen in Anlehnung an Kindling und Schirnbacher (2013), Kindling et al. (2013a) und Burger et al. (2013).

0 Vorwort

Die Humboldt-Universität zu Berlin hat im Spätsommer 2012 damit begonnen, ein institutionelles Konzept zum Forschungsdatenmanagement zu erarbeiten. In diesem Rahmen wurde im 1. Quartal 2013 eine breit angelegte Umfrage durchgeführt, die den aktuellen Stand im Umgang mit Forschungsdaten und die Anforderungen an zentrale Serviceleistungen ermitteln sollte. Der vorliegende Bericht stellt die Umfrage und ihre Ergebnisse dar. Es sollen zudem die Hintergründe der verwendeten Fragen erklärt und damit die Interpretation der Ergebnisse erleichtert werden.

Bei der Gestaltung und Durchführung der Umfrage war es möglich, auf die Erfahrungen mit vergleichbaren Umfragen an anderen Hochschulen zurückzugreifen. Vor allem die Expertise des britischen Digital Curation Centre (DCC) und der von ihm frei zur Verfügung gestellten Materialien waren dabei von besonderer Bedeutung. Außerdem sei an dieser Stelle allen Teilnehmern des Projektseminars am Institut für Bibliotheks- und Informationswissenschaft (IBI) der Humboldt-Universität zu Berlin im Wintersemester 2012/13 für ihr herausragendes Engagement herzlich gedankt.

Einhergehend mit dem Open-Access-Gedanken werden der vorliegende Bericht und die ihm zugrundeliegenden Forschungsdaten öffentlich zugänglich gemacht. Eine aktive Nutzung der Ergebnisse und ein reger Austausch mit anderen Hochschulen und wissenschaftlichen Einrichtungen wird von den Autoren ausdrücklich begrüßt.

1 Einleitung

1.0 Motivation und Hintergrund

Forschungsdaten sind ein wesentlicher Bestandteil der Forschungsergebnisse. Sie unterstützen die gewonnenen Erkenntnisse und eröffnen zahlreiche Nachnutzungspotentiale. In Anlehnung an die Empfehlungen der Deutschen Forschungsgemeinschaft (DFG) hat die Humboldt-Universität zu Berlin einen Grundsatz beschlossen, dem zufolge alle den Veröffentlichungen zugrundeliegenden *Primärdaten* für zehn Jahre zugänglich bleiben sollen.⁵ Dieses als „Grundsätze guter wissenschaftlicher Praxis“ bekannte Dokument der DFG ist für viele wissenschaftliche Einrichtungen in Deutschland maßgeblich.

Die Allianz der deutschen Wissenschaftsorganisationen unterstützt zudem in Übereinstimmung mit wichtigen internationalen Organisationen einen „grundsätzlich offenen Zugang zu Daten aus öffentlich geförderter Forschung“.⁶ Dazu werden Forschungsdaten-Repositoryn aufgebaut oder neue Formen des wissenschaftlichen Publizierens (z.B. *enhanced publications*) gefördert, die eine konventionelle Publikation mit den dazugehörigen Materialien und Informationen anreichern.

Im vorliegenden Bericht wird die Situation einer multidisziplinären Hochschule beschrieben und ausgewertet. Die Grundlage dafür bildet eine Bestandsaufnahme der vorhandenen Forschungsdaten, ihrer Eigenschaften, der aktuellen Praxis in der Archivierung und Zugänglichmachung von Forschungsdaten und der Anforderungen an zukünftige Serviceleistungen.

5 DFG (1998) und Humboldt-Universität zu Berlin (2002)

6 Allianz (2010)

1.1 Forschungsdatenmanagement an Hochschulen

Im Vergleich zu den englischsprachigen Ländern Australien, Großbritannien und den Vereinigten Staaten lassen sich im deutschsprachigen Raum vergleichsweise wenige hochschulweite Initiativen zum Forschungsdatenmanagement finden.⁷ Als erste Ansätze zur umfassenden Ermittlung der Formen von Forschungsdaten und des Umgangs mit ihnen können die Online-Befragungen an der Universität Göttingen und an der Eidgenössischen Technischen Hochschule (ETH Zürich) im Mai und Juni 2007 gewertet werden. Die Befragung an der Universität Göttingen wurde von „nestor“, dem Kompetenznetzwerk zur digitalen Langzeitarchivierung, durchgeführt und im Rahmen des von der DFG geförderten Projektes "Kooperative Langzeitarchivierung für Wissenschaftsstandorte" (KoLaWiss) im Arbeitspaket Datenkategorisierung evaluiert.⁸ An der ETH Zürich wurde im Jahr 2011 eine weitere Online-Umfrage zur Analyse des Ist-Zustands und der Anforderungen an künftige Dienstleistungsangebote durchgeführt.⁹ Darauf basierend wird derzeit eine technische Lösung zur Langzeitarchivierung und die Registrierung von Digital Object Identifiers (DOI) zur dauerhaften Zitierbarkeit von Forschungsdaten entwickelt.¹⁰

In Österreich wurde eine Umfrage zum aktuellen Stand der Forschungsdaten am erst im Jahr 2009 gegründeten Institute of Science and Technology (IST Austria) durchgeführt.¹¹ Durch die Ermittlung der Typen und Eigenschaften der am IST Austria vorhandenen Forschungsdaten wurde die Auswahl einer geeigneten Software und die Entwicklung von zusätzlichen Mehrwertdiensten unterstützt. Als Bausteine des zukünftigen Forschungsdaten-Repository sind u.a. die Dokumentation der Forschungsdaten mit Metadaten und eindeutige Identifizierung mittels DOI-Registrierung vorgesehen.¹²

Ein Beispiel an einer außeruniversitären Wissenschaftsorganisation ist die Forschungsdatenumfrage an den Instituten der Leibniz-Gemeinschaft (WGL). In zwei Erhebungsphasen im Sommer 2010 und Frühjahr 2011 wurde die durch den Arbeitskreis Forschungsdaten der WGL entwickelte Umfrage durch GESIS – Leibniz-Institut für Sozialwissenschaften durchgeführt. Die Befragung erfolgte an 89 Instituten der WGL. Dabei wurden beispielsweise Aspekte wie das Vorhandensein von Metadatenstandards, Data Policies und Forschungsdaten-Repositoryn abgefragt.¹³ Wenngleich zum Zeitpunkt der Verfassung des vorliegenden Berichts die vollständige Auswertung der Ergebnisse an der WGL noch nicht veröffentlicht wurde, kann ein Vergleich der Ergebnisse an universitären und außeruniversitären Einrichtungen eine weitere interessante Forschungsfrage darstellen und zum umfassenderen Verständnis zum Umgang mit Forschungsdaten im deutschsprachigen Raum beitragen.

Zur Unterstützung der Forschenden werden an deutschen Hochschulen derzeit diverse Serviceangebote entwickelt. Dies geschieht oft in fachspezifischen Pilotprojekten wie z.B. in dem auf sozialwissenschaftliche Forschungsdaten ausgerichteten Projekt "Informium" an der Universität Bielefeld. Ferner wird eine regionale Vernetzung zum Aufbau institutionenübergreifenden Lösungen angestrebt, wie z.B. am Wissenschaftsstandort Kiel unter Beteiligung der Christian-Albrechts-Universität zu Kiel. Außerdem werden institutionelle fachübergreifende Repositoryn zur Archivierung und Publikation digitaler Forschungsdaten wie z.B. an der Technischen Universität Berlin entwickelt.¹⁴

Weitere vergleichbare Umfragen zur Ermittlung des aktuellen Stands im Umgang mit

7 Für einen internationalen Überblick zu Forschungsdatenmanagement an Hochschulen s. Burger et al. (2013)

8 Vgl. Vlaeminck (2008)

9 Vgl. Töwe, Scheid (2011)

10 Vgl. ETH Zürich, Projekt Digitaler Datenerhalt <http://www.library.ethz.ch/Ueber-uns/Projekte/Digitaler-Datenerhalt>

11 Vgl. Porsche (2012)

12 Vgl. Porsche (2013)

13 Vgl. Quandt (2012)

14 Vgl. Burger et al. (2013) und Kindling et al. (2013a)

Forschungsdaten bzw. der Anforderungen an unterstützende Serviceangebote an einer wissenschaftlichen Einrichtung im deutschsprachigen Raum sind den Autoren des vorliegenden Berichts nicht bekannt.

1.2 Ausgangslage der Humboldt-Universität zu Berlin

Die Humboldt-Universität zu Berlin wurde nach dem Bildungsideal Wilhelm von Humboldts vor mehr als 200 Jahren gegründet. „Wer hier arbeitet“ – heißt es im Leitbild der Universität – „soll den Willen und die Fähigkeit haben durch eigene Leistungen das Wissen zu mehren und seine verantwortungsvolle Nutzung zu sichern.“¹⁵ Die Idee eines freien, gleichberechtigten Zugangs zum Wissen wurde durch die Verabschiedung der Open-Access-Erklärung im Jahr 2006 nochmals bekräftigt.¹⁶ Zur Umsetzung dieser Idee dient beispielsweise der universitätseigene Dokumenten- und Publikationsserver (edoc-Server).¹⁷

Die Zugänglichmachung von Forschungsdaten an der Humboldt-Universität zu Berlin wurde bisher allerdings nicht ganzheitlich angegangen. Vielmehr finden sich eigenständige Initiativen an Instituten und wissenschaftlichen Einrichtungen der Universität, wie beispielsweise die Geodateninfrastruktur¹⁸ am Geographischen Institut oder das Research Data Center des Sonderforschungsbereichs 649 "Ökonomisches Risiko"¹⁹ an der Wirtschaftswissenschaftlichen Fakultät.

Ein erster Ansatz eines koordinierten universitätsweiten Vorgehens ist die "Satzung über die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis und über den Umgang mit Vorwürfen wissenschaftlichen Fehlverhaltens". Diese im Jahr 2002 vom Akademischen Senat der Universität beschlossenen Grundsätze beziehen sich zum großen Teil auf die "Empfehlungen zur Sicherung guter wissenschaftlicher Praxis" der DFG vom 1998 und fordern die Wissenschaftler der Universität auf, die Primärdaten als Grundlagen für Veröffentlichungen auf haltbaren und gesicherten Trägern in der Arbeitsgruppe oder Einrichtung, in der sie entstanden sind, für zehn Jahre zugänglich zu halten.²⁰

Von der Medienkommission des Akademischen Senats wurde der Umgang mit Forschungsdaten zum ersten Mal in ihrem Multimedia-Förderprogramm für das Jahr 2012 thematisiert. Unter dem Schwerpunkt "Systematischer Aufbau von nachhaltigen Strukturen zur Erfassung, Erschließung und Bereitstellung von Forschungsprimärdaten" konnte die Förderung für Gemeinschaftsprojekte von Instituten und Zentraleinrichtungen zum Aufbau einer neuen Infrastruktur beantragt werden.²¹ Die Förderung derartiger Vorhaben wurde erneut in der darauffolgenden Ausschreibung für das Jahr 2013/14 ermöglicht.²²

15 Leitbild der Humboldt-Universität zu Berlin, beschlossen durch das Konzil der Humboldt-Universität zu Berlin am 13. Februar 2002. Online zugänglich unter <https://www.hu-berlin.de/ueberblick/humboldt-universitaet-zu-berlin/leitbild> zuletzt geprüft am 14.06.2013.

16 Vgl. Open-Access-Erklärung der Humboldt-Universität zu Berlin, verabschiedet am 9. Mai 2006 vom Akademischen Senat der Humboldt-Universität zu Berlin. Online zugänglich unter http://edoc.hu-berlin.de/e_info/oa-erklaerung.php zuletzt geprüft am 14.06.2013.

17 Leitlinien des Dokumenten- und Publikationsservers der Humboldt-Universität zu Berlin vom 1. September 2010. Online zugänglich unter http://edoc.hu-berlin.de/e_info/leitlinien.php zuletzt geprüft am 14.06.2013.

18 <http://gdi.geo.hu-berlin.de/>

19 <http://sfb649.wiwi.hu-berlin.de/>

20 s. § 5 Sicherung und Aufbewahrung von Primärdaten, in: Humboldt-Universität zu Berlin (2002)

21 Förderprogramm 2012 für digitale Medien in Forschung, Lehre und Studium (Multimedia-Förderprogramm 2012). Ausschreibung der Medienkommission des Akademischen Senats. Online zugänglich unter http://gremien.hu-berlin.de/mk/archiv/mm-f12/mm-f12_ausschreibung.pdf zuletzt geprüft am 17.06.2013.

22 Förderprogramm 2013/14 für digitale Medien in Forschung, Lehre und Studium. Ausschreibung der Medienkommission des Akademischen Senats. Online zugänglich unter http://gremien.hu-berlin.de/mk/archiv/mm-f13/mm-f13_ausschreibung.pdf zuletzt geprüft am 17.06.2013.

1.2 Forschungsdatenmanagement-Initiative

Unter Federführung des Vizepräsidenten für Forschung und unter Mitwirkung der Zentraleinrichtungen Computer- und Medienservice, Universitätsbibliothek und Servicezentrum Forschung wurde im Jahr 2012 eine dezidierte Stelle zum Forschungsdatenmanagement gegründet und besetzt. Die Initiative hat es sich zum Ziel gesetzt, ein universitätsweites Konzept zum Umgang mit digitalen Forschungsdaten zu entwickeln. Damit verbundene Aktivitäten werden durch den Lehr- und Forschungsbereich Informationsmanagement am IBI fachlich begleitet.²³ Eine der ersten Aufgaben der eingerichteten Stelle war die Öffentlichkeitsarbeit: Neben einer speziellen Webseite²⁴ mit den wichtigsten Informationen zum Thema Forschungsdaten wurde ein einführender Artikel in der Universitätszeitung²⁵ veröffentlicht. Weitere Arbeitsbereiche waren die Bestandsaufnahme des Ist-Zustands im Umgang mit Forschungsdaten an der HU mittels der vorliegenden Umfrage und die Analyse der Vorgehensweisen zur Organisation des Forschungsdatenmanagements an anderen Hochschulen weltweit. Letzteres erfolgte im Rahmen eines Projektseminars am IBI zusammen mit den Studierenden im Wintersemester 2012/13. Als Ergebnis des Projektseminars ist ein ausführlicher Bericht entstanden, der auf Grundlage der Erfahrungen an anderen Hochschulen mögliche Szenarien für weiteres Vorgehen an der HU vorstellt.²⁶ Weitere Schritte sind derzeit in Vorbereitung und werden über die Webseite umgehend bekanntgegeben.²⁷

2 Methodik

Im folgenden Kapitel wird die Gestaltung und Durchführung der Umfrage näher beschrieben.

2.1 Gestaltung des Fragebogens

Ausgehend von dem Ziel der Umfrage, den aktuellen Stand im Umgang mit digitalen Forschungsdaten und die Anforderungen an zukünftige Serviceangebote an der Humboldt-Universität zu ermitteln, richtete sich das Vorhaben vorwiegend an das wissenschaftliche Personal der Universität. Diese Abgrenzung erfolgte unter der Annahme, dass genau diese Zielgruppe in ihrer wissenschaftlichen Tätigkeit die Forschungsdaten produziert bzw. bearbeitet.

Bei der Gestaltung des Fragebogens war es möglich, sich auf die Erfahrungen mit vergleichbaren Umfragen an anderen Hochschulen zu stützen. Insbesondere konnten in diesem Zusammenhang die Fragebögen der University of Glasgow und des Imperial College London (Projekt "Data Asset Framework"), der University of Cambridge (Projekt "Incremental") und der ETH Zürich (aufbauend auf der Umfrage der Leibniz-Gemeinschaft) nachgenutzt werden. Außerdem waren die Erkenntnisse aus dem PARSE.Insight-Projekt ("Permanent Access to the Records of Science in Europe") im Siebten Rahmenprogramm der Europäischen Union hilfreich.²⁸

In Anlehnung an o.g. Hilfsmittel ist ein 24 Fragen umfassender Fragebogen in zwei Sprachen (Deutsch und Englisch) entstanden. Zur besseren Handhabbarkeit wurde dieser in fünf Fragengruppen strukturiert: Fachbereich und Position des Teilnehmers, Datentypen (Eigenschaften und Speicherung von Forschungsdaten), Zugang (Veröffentlichung und Nachnutzung von Forschungsdaten, Richtlinien), Infrastruktur und Service (vorhandene und gewünschte Infrastruktur, hilfreiche Kriterien) und weiteres Interesse (Kontaktangaben bei Interesse an einem persönlichen Interview oder Umfrageergebnissen). Der Fragebogen wurde zunächst an die Ausgangslage der

23 Vgl. Kindling et al. (2013a)

24 <http://www.cms.hu-berlin.de/dataman>

25 Simukovic (2012)

26 S. Burger et al. (2013)

27 Für einen zusammenfassenden Artikel des vorliegenden Berichts s. Kindling et al. (2013a)

28 Für genaue Angaben s. Literaturverzeichnis

Humboldt-Universität angepasst und anschließend in Diskussionen mit den Beteiligten an der Initiative und Studierenden im Projektseminar weiterentwickelt. Zudem folgte eine Testphase mit einem kleinen Teil der Zielgruppe, deren Rückmeldungen ebenso zur Verbesserung beigetragen haben. Die endgültige Version des Fragebogens (in deutscher Sprache) steht am Ende des Berichts als Anhang 7.2 zur Verfügung.

2.2 Durchführung der Umfrage

Um die Vorteile der digitalen Kommunikation in der aktuellen Umfrage zu nutzen, wurde mithilfe des an der HU bereits eingesetzten Open-Source-Produktes "LimeSurvey"²⁹ ein Online-Fragebogen erstellt. Die Teilnahme an der Umfrage wurde in zwei Phasen organisiert: Zuerst wurde die Einladung in Kooperation mit dem Servicezentrum Forschung an die Leiter der Forschungsprojekte versandt. Drei Wochen später folgte eine Einladung an das gesamte wissenschaftliche Personal der Universität. Um die gesamte Zielgruppe zu erreichen, wurde zudem ein spezieller E-Mail-Verteiler erstellt, der die wissenschaftlichen Mitarbeiter (Mittelbau) und die Professuren umfasste.³⁰ Insgesamt konnte der Fragebogen vom 24. Januar bis 8. März 2013 ausgefüllt werden.

Während der Laufzeit der Umfrage wurde diese intensiv (durch die Verantwortlichen der FDM-Initiative) begleitet. Dies umfasste insbesondere die Anpassung der Hilfetexte an problematischen Stellen³¹ sowie die zusätzliche gezielte Kontaktaufnahme mit weniger aktiven Instituten und wissenschaftlichen Einrichtungen der Universität. Diese Vorgehensweise ermöglichte es, den Rücklauf der Umfrage nicht nur zu erhöhen, sondern auch eine breite Verteilung der Antworten unter den Instituten und wissenschaftlichen Einrichtungen zu erreichen. So können die Ergebnisse der Umfrage an der Humboldt-Universität sowohl als Ganzes betrachtet werden, als auch im Vergleich der Antworten zwischen verschiedenen Teilnehmergruppen. Außerdem hat sich die bei fast allen Fragen angebotene Möglichkeit, die Antworten mit eigenen Freitext-Kommentaren zu vervollständigen, als sehr nützlich erwiesen. Wir erhielten somit über 600 Kommentare, die sehr wertvolle Einsichten in Arbeitsweisen verschiedener Fachbereiche und den Berufsalltag einzelner Wissenschaftler liefern.³²

Darüber hinaus wurde am Ende des Fragebogens eine Möglichkeit angeboten, bei Interesse an den Ergebnissen oder an einem persönlichen Interview die Kontaktdaten einzutragen (s. Kapitel 3.5). Auf dieser Weise wurde die Gelegenheit geschaffen, mit interessierten Teilnehmern einen direkten Kontakt zu knüpfen und tiefere Einblicke in fachspezifische Arbeitsweisen und Anforderungen an Unterstützungsangebote zu gewinnen. Abgesehen davon wurden keine weiteren personenbezogenen oder -bezieharen Daten erfasst. Damit konnte sich jeder Teilnehmer frei dafür entscheiden, bei der Beantwortung der Fragen anonym zu bleiben.³³

29 <https://umfrage.hu-berlin.de/>

30 Vgl. Personalstatistik der HU <https://www2.hu-berlin.de/personalstatistik/components/personal/daten.php>

31 Bei den meisten Fragen wurden zusätzliche Freitext-Kommentare erlaubt, so dass unklare Stellen umgehend identifiziert und deutlicher formuliert werden konnten.

32 Für eine vollständige Liste mit allen Kommentaren s. Simukovic et al. (2013a)

33 Die Durchführung der Umfrage wurde zusätzlich mit dem Datenschutzbeauftragten der Humboldt-Universität zu Berlin abgestimmt.

3 Ergebnisse

Die Ergebnisse der Umfrage werden im Folgenden gemäß der ursprünglichen Gliederung in Fragengruppen präsentiert. Diesem Teil des Berichts liegt die Auswertung der Ergebnisse in dem Tabellendokument „Umfrage_Forschungsdaten_2013_HU-Berlin.ods“ zugrunde.³⁴

3.1 Teilnehmer und Rücklauf

Der Rücklauf der Umfrage lag mit insgesamt 499 Teilnehmern bzw. 24 % deutlich über dem an Rücklaufquoten an anderen Hochschulen gemessenen Erwartungswert.³⁵ Der Fragebogen wurde überwiegend in deutscher Sprache beantwortet (475 von 499 Teilnehmern bzw. 95 %). Bei der Gestaltung des Fragebogens wurde eine Gesamtzeit von 10 Minuten für die Beantwortung aller Fragen angestrebt. Die erfasste durchschnittliche Beantwortungsdauer lag mit etwa 10,5 Minuten bzw. einem Medianwert von etwa 7,5 Minuten ziemlich genau in diesem Bereich.³⁶ Damit kann der Umfang des Fragebogens für die geplante Dauer als angemessen angesehen werden.

3.1.1 Einrichtungen

Die Umfrage richtete sich an Forschende aus allen Instituten und wissenschaftlichen Einrichtungen der Universität. Die Verteilung der Teilnehmer nach Einrichtungen wird in der Abbildung 1 dargestellt:

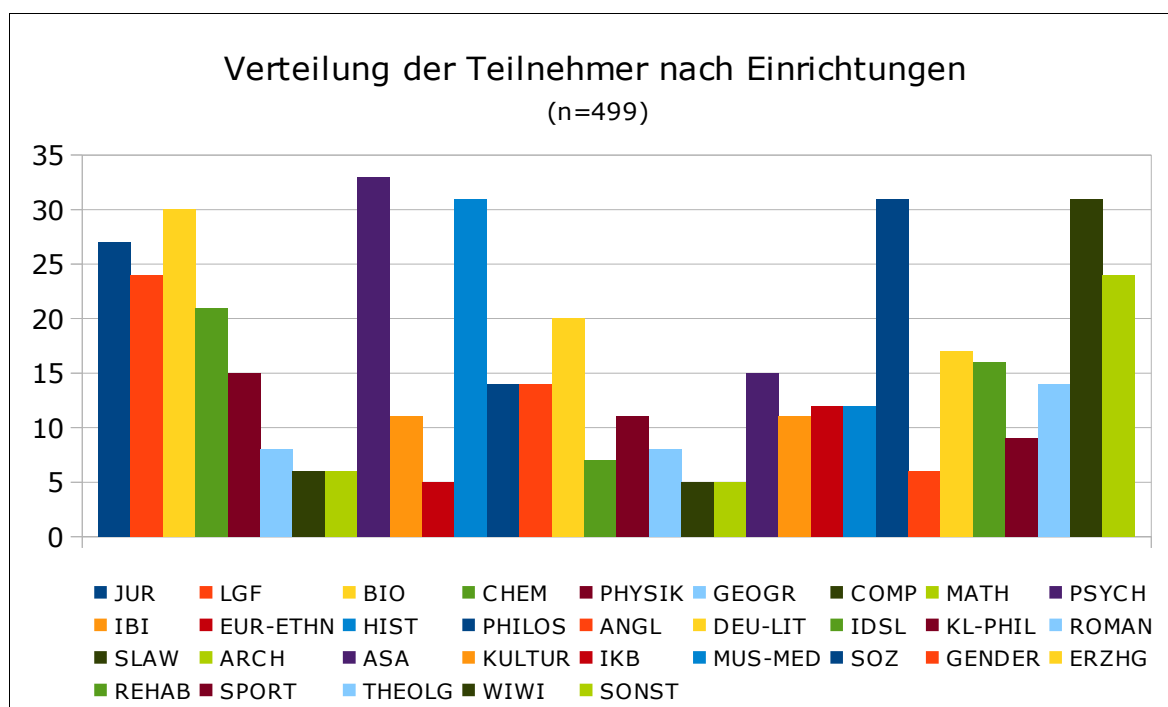


Abbildung 1: Verteilung der Teilnehmer nach Einrichtungen

Die höchste Zahl der Teilnehmer haben das Institut für Psychologie ('PSYCH'; 33), das Institut für Geschichtswissenschaften ('HIST'; 31), das Institut für Sozialwissenschaften ('SOZ'; 31), die

34 Eine komprimierte Darstellung von einigen Umfrageergebnissen findet sich in teilweise gleichem Wortlaut auch in Kindling et al. (2013a). Der vorliegende Bericht lag vor dem Artikel bereits in einer HU-internen Version vor, weshalb der Artikel auf diesen Bericht verweist, nicht aber umgekehrt.

35 Bei den vergleichbaren Umfragen an anderen Hochschulen lag der Rücklauf oft bei 310-370 Teilnehmern bzw. 13-17 % der Zielgruppe.

36 446 von 499 Teilnehmern haben den Fragebogen vollständig ausgefüllt.

Wirtschaftswissenschaftliche Fakultät ('WIWI'; 31), das Institut für Biologie ('BIO'; 30) und die Juristische Fakultät ('JUR'; 27) erreicht. Die Gliederung nach Einrichtungen entspricht dabei der Organisationsstruktur der Universität. Wichtig ist an dieser Stelle anzumerken, dass die Größe des jeweiligen Instituts bzw. der (Mono)Fakultät nicht berücksichtigt wird und die Abbildung sich nur auf die absolute Teilnehmerzahl bezieht.³⁷

3.1.2 Teilnehmergruppen

Im Hinblick auf verschiedene Teilnehmergruppen wird in der Abbildung 2 ersichtlich, dass die Ergebnisse der Umfrage stark von der Gruppe der wissenschaftlichen Mitarbeiter geprägt werden:

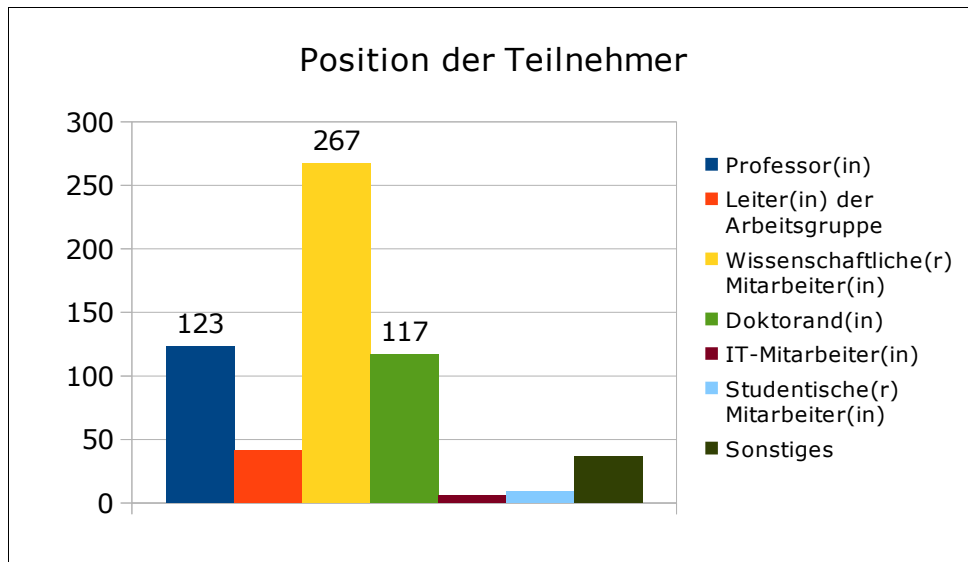


Abbildung 2: Teilnehmergruppen

Darunter haben 63 Teilnehmer als „Wissenschaftliche(r) Mitarbeiter(in)“ und „Doktorand(in)“³⁸ geantwortet (Mehrfachantworten erlaubt). Gemessen an der "potentiellen Teilnehmerzahl" je Gruppe zeigte sich jedoch eine unterschiedliche Dynamik:

Abgleich mit HU-Personalstatistik	Personen	Teilnehmerzahl	Beteiligung an der Umfrage
Professuren	426	123	28,87%
Mittelbau	2015	267	13,25%

Tabelle 1: Abgleich der Teilnehmergruppen mit der HU-Personalstatistik³⁹

Aus dem Abgleich mit der Personalstatistik der HU geht hervor, dass die Professoren etwa doppelt so aktiv wie die wissenschaftlichen Mitarbeiter waren. Eine insgesamt höhere Aktivität seitens der Professoren könnte u.a. dadurch erklärt werden, dass bei der Umfrage verstärkt die Projektleiter angesprochen wurden.⁴⁰ Da die Verantwortung für die Leitung eines Projektes oft mit der hierarchischen Struktur verbunden ist, ist zu vermuten, dass der Anteil der Professoren in dieser

37 Für eine Gewichtung unter Berücksichtigung der Größe der jeweiligen Einrichtung notwendige Personalstatistik ist nicht vorhanden.

38 Einige Umfrageergebnisse wurden exemplarisch im Hinblick auf die Antworten der Doktoranden ausgewertet. S. Kindling (2013c).

39 Gezählt werden Personen ("Köpfe"), Haushalts- und Drittmittel; Stichtag 31.01.2013. Quelle und nähere Erläuterungen: <https://www2.hu-berlin.de/personalstatistik/components/home/start.php>

40 Einerseits wurden die Projektleiter bereits in der ersten Phase eingeladen, andererseits wurde die Einladung an das gesamte wissenschaftliche Personal drei Wochen später an diese Gruppe zum zweiten Mal zugestellt

Gruppe höher ist. Andererseits war die Beteiligung unter den Instituten und wissenschaftlichen Einrichtungen sehr unterschiedlich, so dass dieses Ergebnis auch durch das Engagement einzelner Personen erklärt werden kann.

3.2 Datentypen

In den folgenden Unterkapiteln werden die Ergebnisse der Fragengruppe „Datentypen“ zu den wesentlichen Eigenschaften von Forschungsdaten vorgestellt.

3.2.1 Quellen

In der Fragengruppe "Datentypen" wurden die Teilnehmer zunächst gebeten, ihre Quellen von Forschungsdaten zu identifizieren. Die Ergebnisse werden in der Abbildung 3 dargestellt:

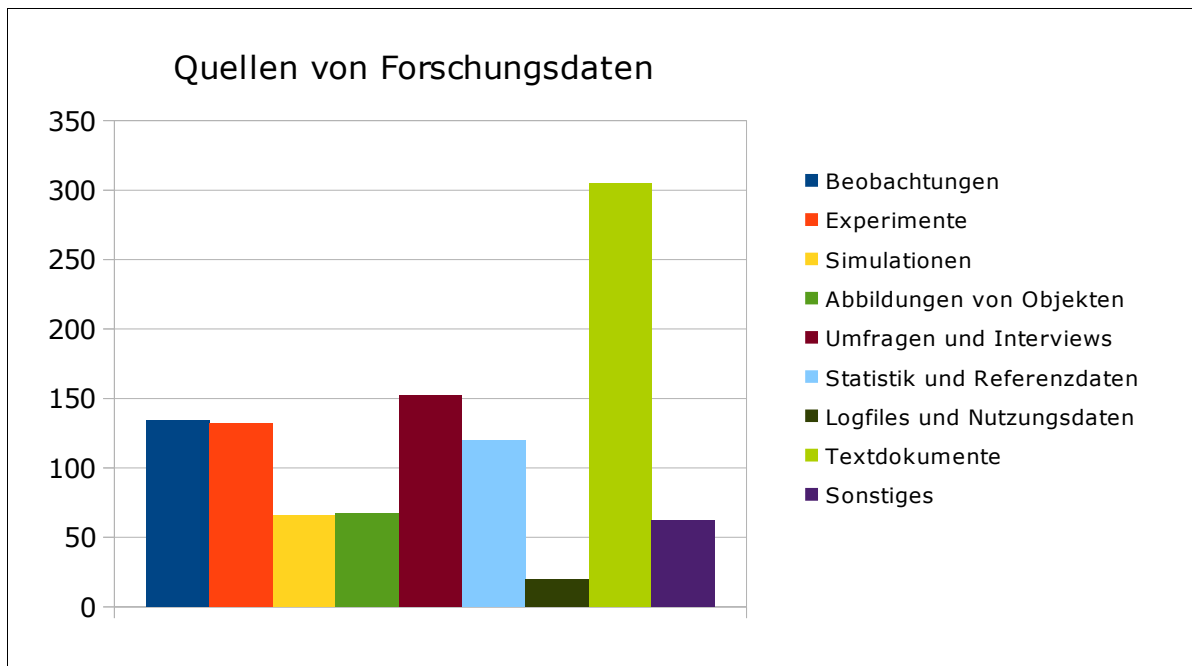


Abbildung 3: Quellen von Forschungsdaten

Mit großem Abstand wurden Textdokumente als häufigste (305 Antworten) Forschungsdatenquellen benannt. Das Ergebnis verwundert nicht, werden doch Textdokumente fachübergreifend als zentrale Form der Vermittlung von Forschungsergebnissen verwendet. Der Reihe nach folgten Umfragen und Interviews (152), Beobachtungen (134), Experimenten (132) und Statistik und Referenzdaten (120).

3.2.2 Typen von Forschungsdaten

Weiterhin sollten die Typen von Forschungsdaten näher spezifiziert werden. Die Antworten zu dieser Frage sollten aufzeigen, welche Datentypen an der HU von hoher Relevanz sind, für die ggf. Infrastrukturlösungen benötigt werden. Die Ergebnisse werden in der Abbildung 4 dargestellt:

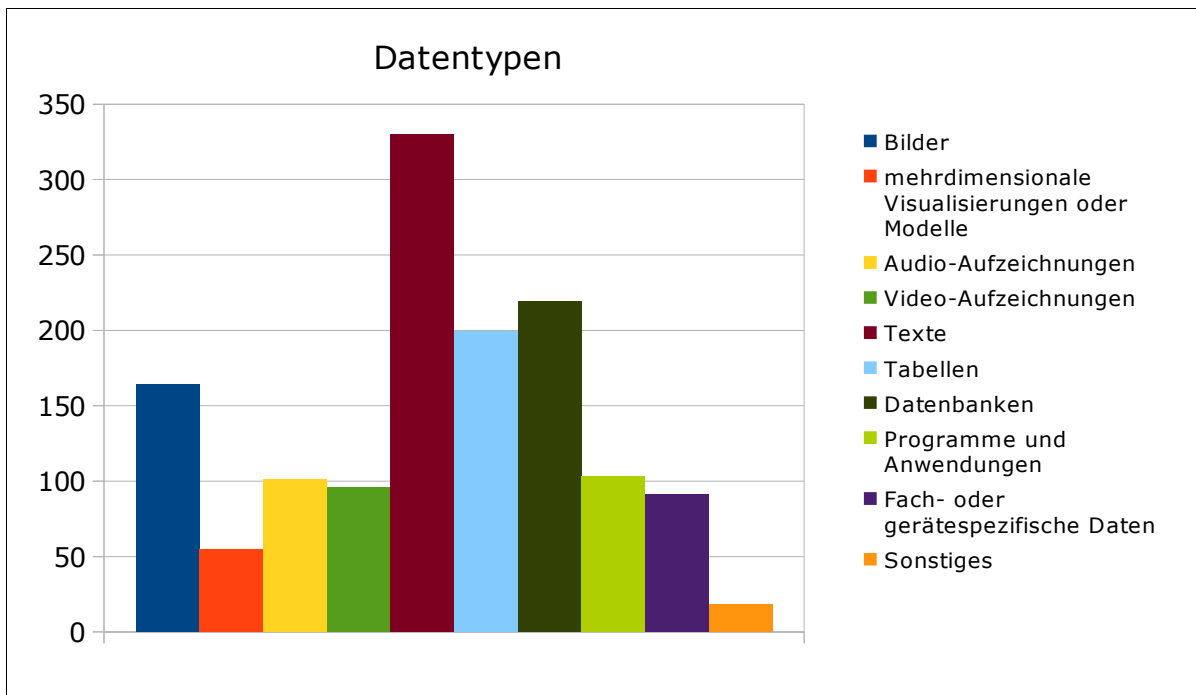


Abbildung 4: Typen von Forschungsdaten

Auch hier wurden Texte am häufigsten benannt (330 Antworten), gefolgt von Datenbanken (219), Tabellen (199) und Bildern (164). Das zeigt unter Berücksichtigung der vorangehenden Resultate, dass verschiedene Datentypen bzw. Anwendungen durchgehend fachübergreifend zu finden sind, auch wenn sich deren genaue Ausprägungen oder Analyseschritte von Fach zu Fach unterscheiden können.

3.2.3 Fach- oder gerätespezifische Daten

Die nächste Frage bot optional die Möglichkeit, fach- oder gerätespezifische Typen von Forschungsdaten genauer zu präzisieren. Die Ergebnisse werden in der Abbildung 5 dargestellt:

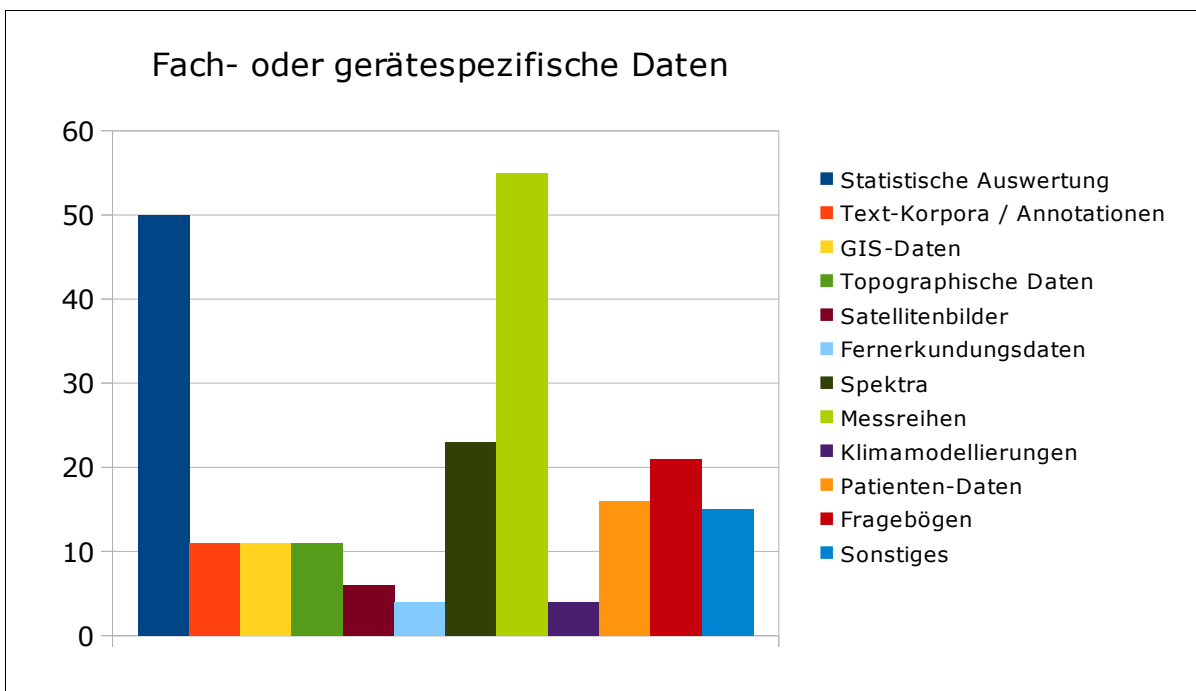


Abbildung 5: Fach- oder gerätespezifische Daten

Unter den angebotenen Antwortoptionen wurden am häufigsten Messreihen (55) und statistische Auswertungen (50) ausgewählt, gefolgt von Spektren (23), Fragebögen (21) und Patienten-Daten (16). Außerdem wurden unter 'Sonstiges' (15) weitere Typen genannt. Insgesamt bieten die Ergebnisse zu den ersten drei Fragen einen guten Einblick in die Arbeitsweisen, mit deren Hilfe die Forschungsdaten an der Humboldt-Universität zu Berlin produziert bzw. bearbeitet werden.

3.2.4 Speicherorte

Die Teilnehmer wurden danach gefragt, wo sie ihre Forschungsdaten speichern. Die Ergebnisse werden in der Abbildung 6 dargestellt:

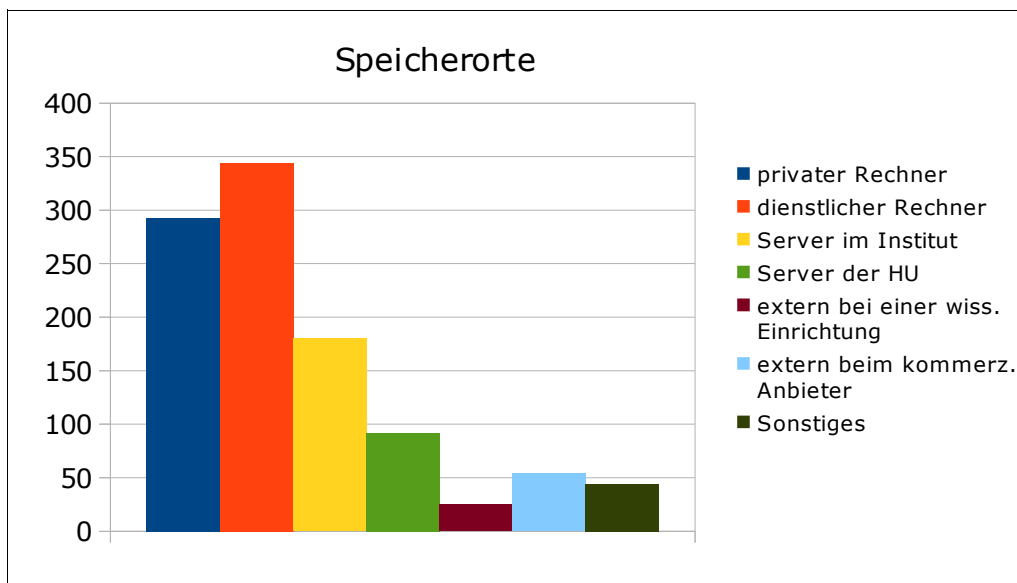


Abbildung 6: Speicherorte

Die Ergebnisse zeigen eine überwiegend lokale Speicherung von Forschungsdaten auf dienstlichen Rechnern (344) und privaten Rechnern (293), gefolgt von der Speicherung auf institutseigenen Servern (180). Die Antworten in Freitext-Kommentaren bestätigten auch die Vermutung, dass Cloud-Dienste wie beispielsweise Dropbox von den Wissenschaftlern an der Humboldt-Universität benutzt werden.

3.2.5 Sicherung auf Datenträgern

Anschließend wurden die Teilnehmer gefragt, auf welchen Datenträgern sie ihre Forschungsdaten zusätzlich sichern. Die Antworten zu dieser und den folgenden Fragen (s. Kapitel 3.2.6-8) sollten ansatzweise einen ersten Einblick in die Sicherungsstrategie einzelner Wissenschaftler geben. Die Ergebnisse werden in der Abbildung 7 dargestellt:

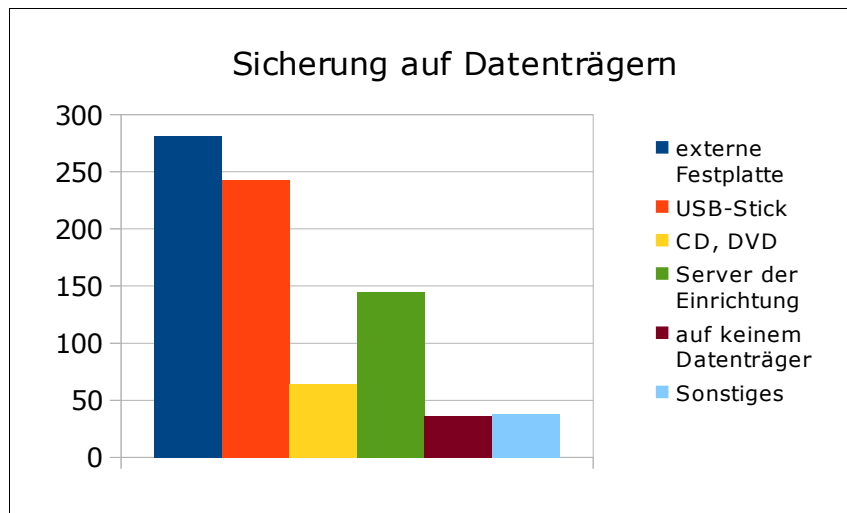


Abbildung 7: Sicherung auf Datenträgern

Wie aus der Abbildung 7 hervorgeht, werden die Forschungsdaten meist auf externen Datenträgern wie Festplatten (281) und USB-Sticks (243) gesichert. Ein geringerer Teil wird zentral auf dem Server der Einrichtung (145) oder bei Cloud-Diensten (19 Antworten unter 'Sonstiges') gesichert. Einige Teilnehmer nutzen dafür auch CDs oder DVDs (64). Nur wenige Teilnehmer (36) führen keine zusätzliche Sicherung der Forschungsdaten durch.

3.2.6 Gesamtgröße der Forschungsdaten

Die Entscheidungen über einen Speicherort oder einen Datenträger können mit der Größe der Forschungsdaten zusammenhängen. Die Teilnehmer wurden daher gebeten, die Gesamtgröße ihrer Forschungsdaten einzuschätzen. Die Antworten werden in der Abbildung 8 dargestellt:

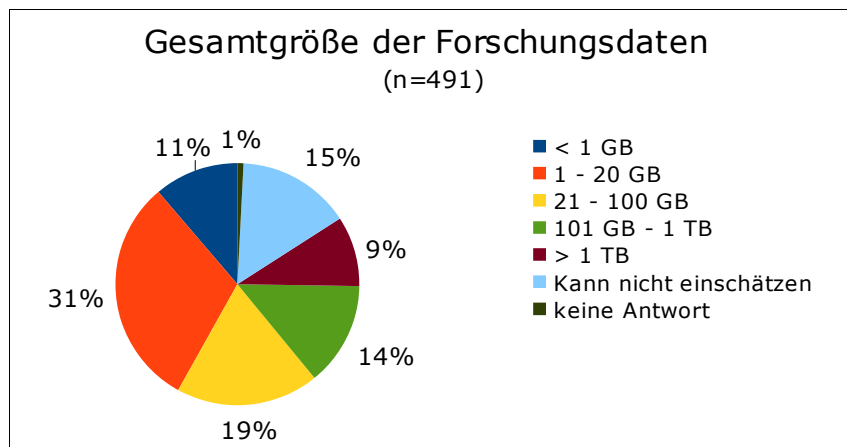


Abbildung 8: Gesamtgröße der Forschungsdaten

Die meisten Teilnehmer (151; 31 %) schätzten die Gesamtgröße ihrer Forschungsdaten auf 1-20 GB, gefolgt von 21-100 GB (93; 19 %) und 101 GB-1 TB (68; 14 %). Ein beträchtlicher Teil der Teilnehmer konnte die Größe nicht einschätzen (74; 15 %). Dies ist möglicherweise auf die Komplexität der Definition von Forschungsdaten bzw. deren genauer Ausprägung im jeweiligen Fachbereich zurückzuführen. Bei näherer Analyse ließ sich erkennen, dass Unsicherheit insbesondere unter Teilnehmern aus den Geisteswissenschaften herrschte, was vermutlich im Zusammenhang mit dem Begriffsverständnis von Forschungsdaten steht. Zusammenfassend kann man sagen, dass die Mehrheit der Teilnehmer sich im Bereich von etwa 20 GB bewegen. Diese Erkenntnis kann u.a. bei der Planung eines Dienstes zur zentralen Speicherung von

Forschungsdaten hilfreich sein.⁴¹

3.2.7 Erstellung der Sicherungskopien

Die Teilnehmer der Umfrage wurden weiterhin gefragt, wie oft Sicherungskopien von Forschungsdaten erstellt werden. Die Ergebnisse werden in der Abbildung 9 dargestellt:

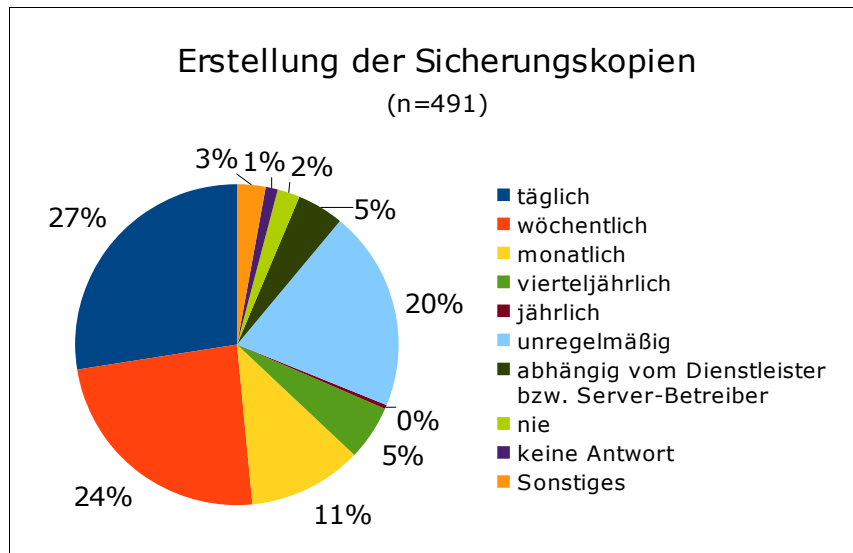


Abbildung 9: Erstellung der Sicherungskopien

Die meisten Teilnehmer gaben an, ihre Forschungsdaten täglich (135; 27 %) oder wöchentlich (118; 24 %) zu sichern. Ein relativ hoher Anteil der Teilnehmer verwies auf eine unregelmäßige Sicherung (99; 20 %). Die Ergebnisse zeigen, dass sich die Praxis von Fall zu Fall unterscheidet.

3.2.8 Verantwortung für Forschungsdaten

Des weiteren wurden die Teilnehmer gefragt, wer für die Speicherung, Sicherung oder Archivierung der Forschungsdaten verantwortlich ist. Die Ergebnisse werden in der Abbildung 10 dargestellt:

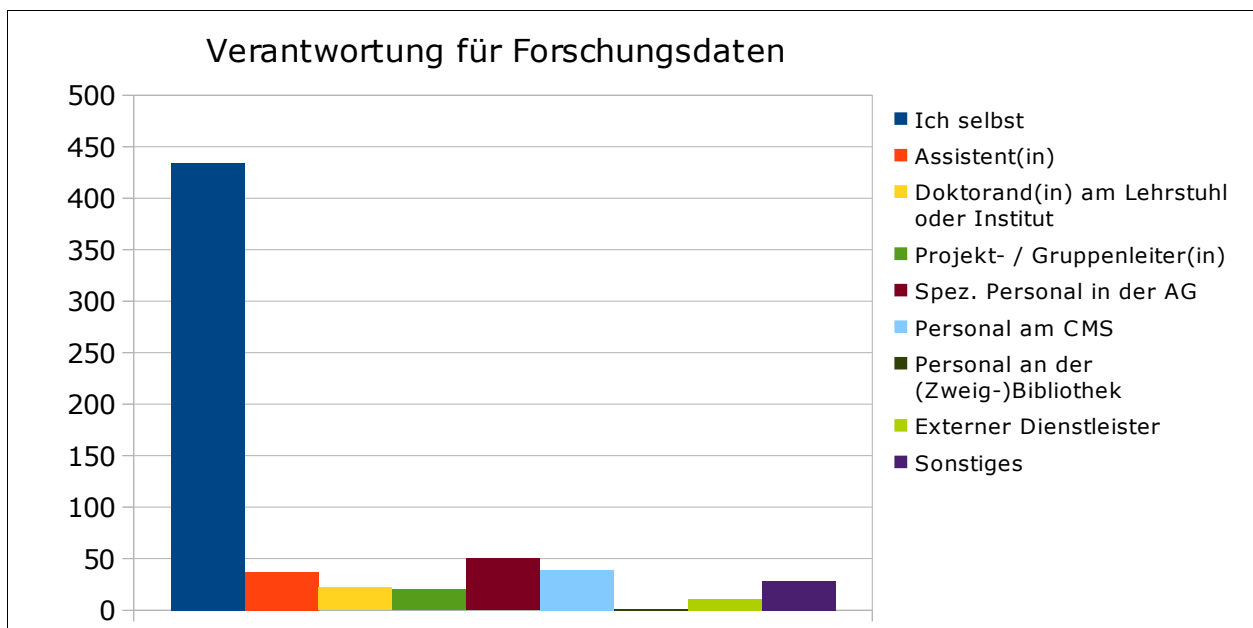


Abbildung 10: Verantwortung für Forschungsdaten

41 Vgl. Kindling et al. (2013a)

Die überwiegende Mehrheit der Teilnehmer ist für die Speicherung, Sicherung oder Archivierung ihrer Forschungsdaten selbst verantwortlich (434). Daneben wird diese Aufgabe teilweise an spezielles Personal in der Arbeitsgruppe (50), Personal am CMS (39) oder an eigene Assistenten (37) delegiert. Bei dieser Frage wurden auch Freitext-Kommentare erlaubt, in denen u.a. Studentische Hilfskraft, EDV-Personal bzw. Server am Institut, Projektpartner und Sekretärin genannt wurden.

3.3 Zugang

In den folgenden Unterkapiteln werden die Ergebnisse der Fragengruppe „Zugang“ zur gängigen Praxis bei der Zugänglichmachung von Forschungsdaten vorgestellt.

3.3.1 Zugriff auf Forschungsdaten

Zunächst wurden die Teilnehmer befragt, wer auf ihre Forschungsdaten zugreifen kann bzw. darf. Die Ergebnisse werden in der Abbildung 11 dargestellt:

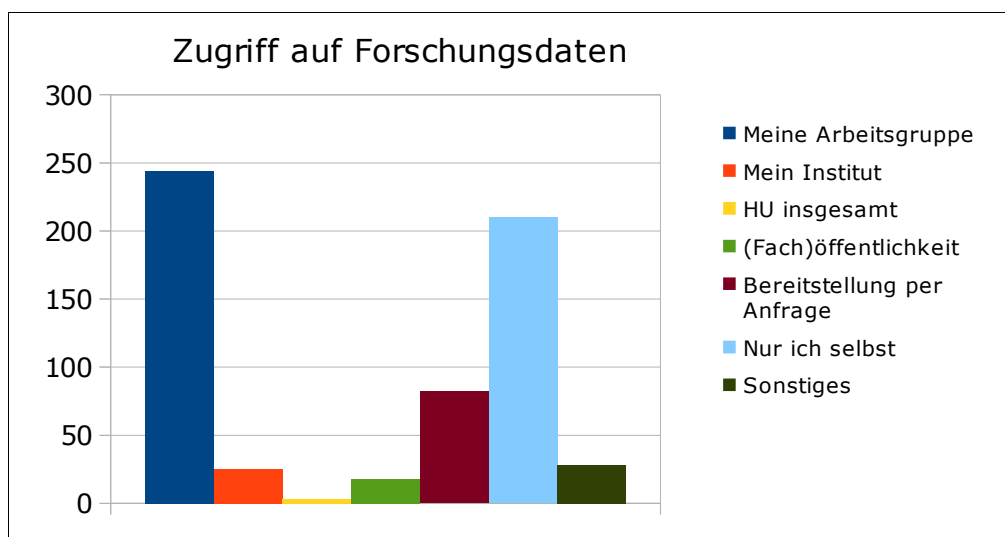


Abbildung 11: Zugriff auf Forschungsdaten

Die meisten Teilnehmer gaben an, dass der Zugriff auf Forschungsdaten ausschließlich auf ihre Arbeitsgruppe (244) oder sie selbst (210) beschränkt ist. Zudem haben manche Teilnehmer in Freitext-Kommentaren präzisiert, dass dies vom konkreten Projekt oder den produzierten Forschungsdaten abhängig ist und der Zugriff nur den Projektbeteiligten (innerhalb oder außerhalb der Arbeitsgruppe an der eigenen Einrichtung) ermöglicht wird. Bei Promotionsvorhaben wurden auch der Doktorvater oder der Betreuer genannt. Außerdem gaben einige wenige Teilnehmer in Freitext-Kommentaren an, dass die Forschungsdaten öffentlich zugänglich gemacht werden.

3.3.2 Veröffentlichung und Nachnutzung

Daran anknüpfend wurden die Teilnehmer bezüglich ihrer Erfahrungen bei der Veröffentlichung und Nachnutzung von Forschungsdaten befragt. Die Ergebnisse werden in folgenden Abbildungen 12a - c dargestellt:

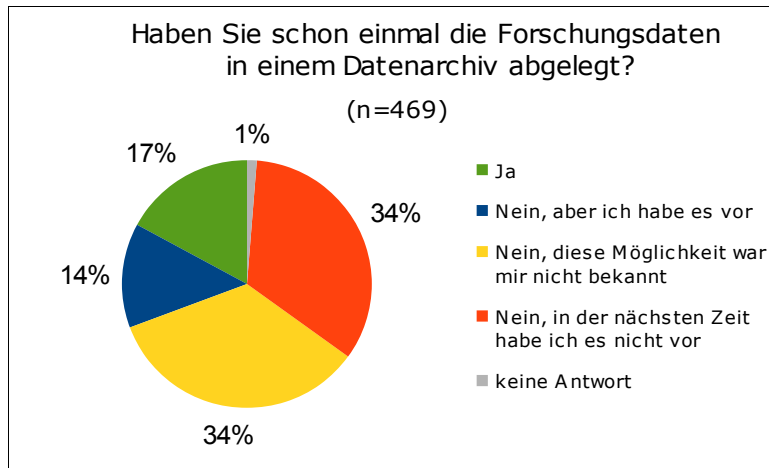


Abbildung 12a: Veröffentlichung und Nachnutzung: In einem Datenarchiv abgelegt?

Die Ergebnisse der ersten Teilfrage zeigen, dass 80 Teilnehmer (17 %) ihre Forschungsdaten mindestens einmal in einem Datenarchiv abgelegt haben und weitere 64 (14 %) dies vorhaben. Den meisten Teilnehmern war diese Möglichkeit allerdings unbekannt (161; 34 %) oder sie haben es nicht vor (158; 34 %).

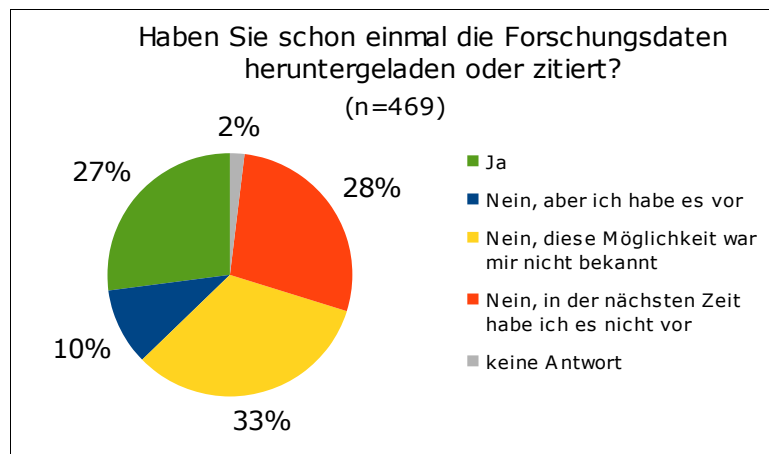


Abbildung 12b: Veröffentlichung und Nachnutzung: Heruntergeladen oder zitiert?

Ähnlich vielen Teilnehmern war die Möglichkeit nicht bekannt, dass Forschungsdaten anderer Forscher aus einem Datenarchiv heruntergeladen oder zitiert werden können (154; 33 %), bzw. haben sie dies nicht vor (131; 28 %). Bei der Nutzung fremder Daten haben deutlich mehr Teilnehmer eigene Erfahrungen gesammelt (127; 27 %) bzw. haben dies vor (48; 10 %).

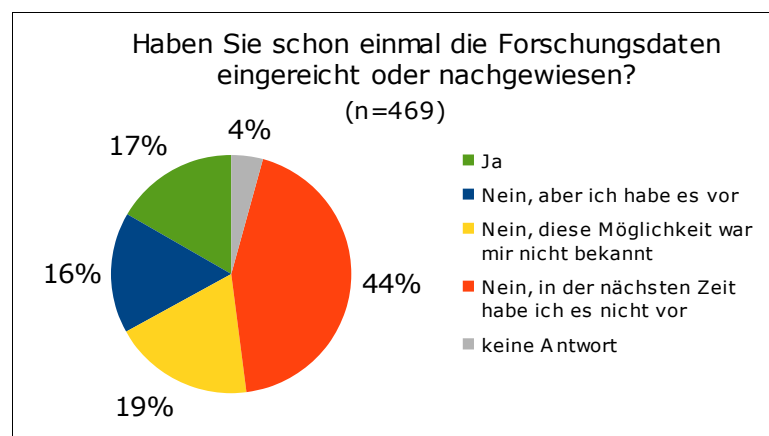


Abbildung 12c: Veröffentlichung und Nachnutzung: Eingereicht oder nachgewiesen?

Bei der dritten Teilfrage wurden die Teilnehmer gefragt, ob sie ihre Forschungsdaten schon einmal bei einer Zeitschrift eingereicht bzw. nachgewiesen haben. Die Ergebnisse zeigen, dass dies bei 78 Teilnehmern (17 %) der Fall war und weitere 77 (16 %) dies vorhaben. Auffallend ist dabei, dass diese Möglichkeit vergleichsweise wenigen Teilnehmern (89; 19 %) unbekannt ist und gleichzeitig eine große Mehrheit der Teilnehmer (205; 44 %) dies dennoch in der nächsten Zeit auch nicht vorhaben. Um genauer sagen zu können, ob diese Tendenz mit allgemeinen Bedenken bzgl. der Zugänglichmachung von Forschungsdaten oder der Position gegenüber konkreten Fachzeitschriften oder Verlagen verbunden ist, muss eine genauere Untersuchung erfolgen (z.B. durch Interviews mit einzelnen Wissenschaftlern).

In der Abbildung 13 werden die Ergebnisse aller drei Teilfragen in einer Matrix visualisiert:

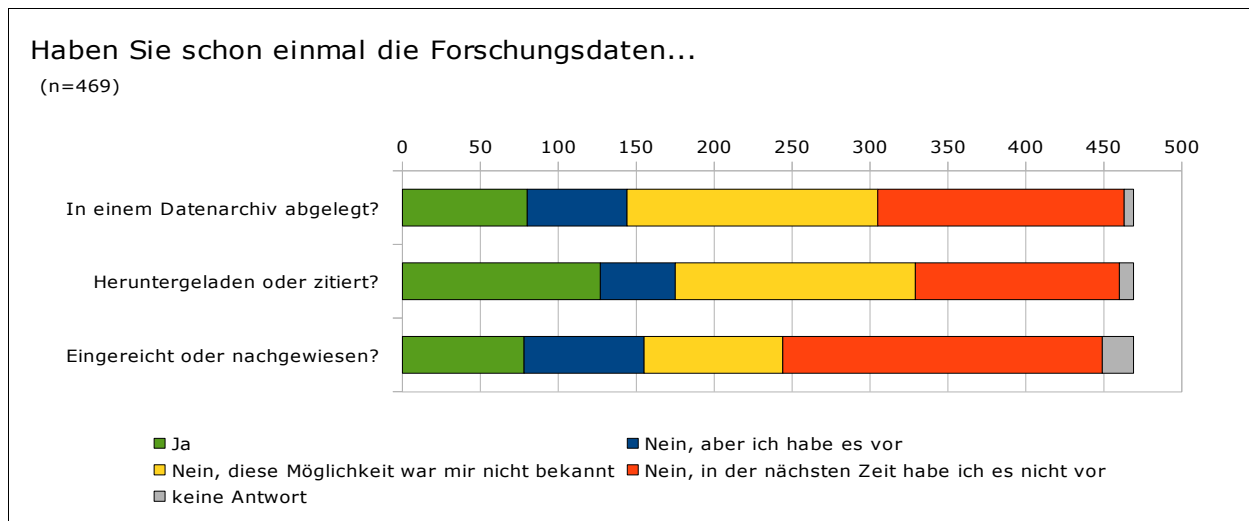


Abbildung 13: Veröffentlichung und Nachnutzung (Matrix)

Diejenigen Teilnehmer, die mit „Ja“ oder „Nein, aber ich habe es vor“ geantwortet haben, wurden zudem gebeten, das betreffende Datenarchiv bzw. die Zeitschrift zu benennen. Daraus folgten 116 Angaben mit teilweise mehrfachen Nennungen, darunter Zeitschriftentitel der Public Library of Science (PLoS), GenBank (DNA-Sequenzdatenbank des National Institutes of Health), Datenbanken des GESIS – Leibniz-Institut für Sozialwissenschaften u.v.m. Eine vollständige Liste mit allen Nennungen steht in dem dem Bericht zugrundeliegenden Tabellendokument zur Verfügung.⁴²

3.3.3 Sicherung guter wissenschaftlicher Praxis

Die folgende Frage hatte zum Ziel, die Bekanntheit und die praktische Umsetzung der Grundsätze zur Sicherung guter wissenschaftlicher Praxis zu klären. Diese Grundsätze wurden vom Akademischen Senat der Universität im Jahr 2002 beschlossen und regeln die Aufbewahrung von wissenschaftlichen Daten⁴³: „Primärdaten als Grundlagen für Veröffentlichungen sollen auf haltbaren und gesicherten Trägern in der Arbeitsgruppe/ Einrichtung, wo sie entstanden sind, für zehn Jahre zugänglich bleiben.“ Die Ergebnisse werden in der Abbildung 14 dargestellt:

42 s. Simukovic et al. (2013a)

43 Die Begriffe „Primärdaten“ und „Forschungsdaten“ werden in diesem Zusammenhang synonym verwendet.

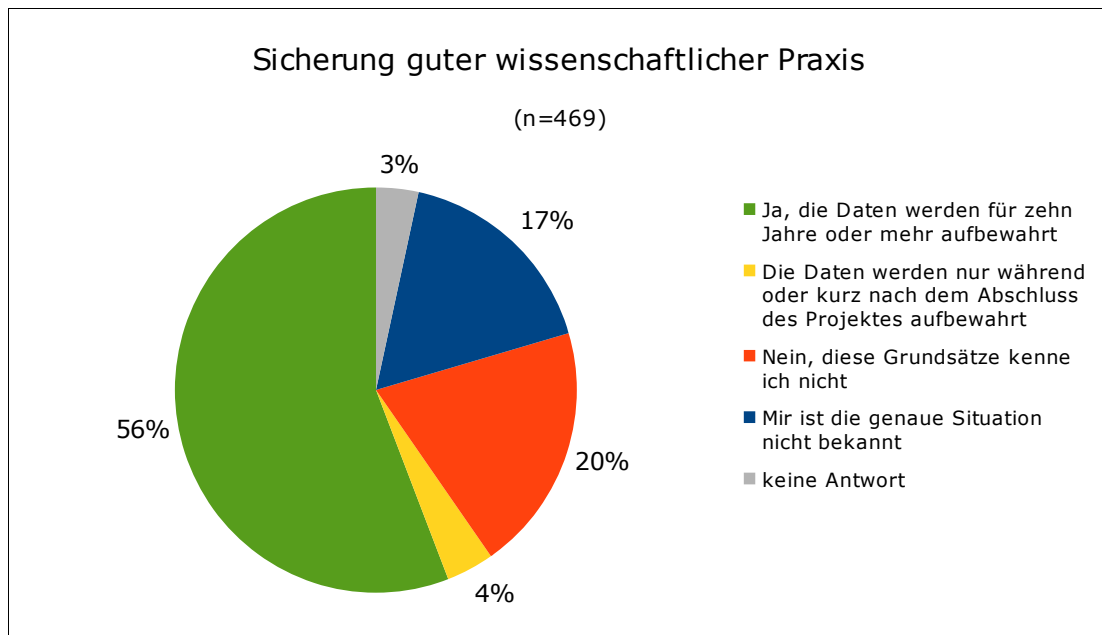


Abbildung 14: Sicherung guter wissenschaftlicher Praxis

Die Mehrzahl der Teilnehmer (262; 56 %) bewahrt die Daten den Ergebnissen zufolge für zehn Jahre oder mehr auf. Ein großer Teil der Teilnehmer war allerdings mit diesen Grundsätzen (93; 20 %) bzw. der genauen Situation zu deren Umsetzung (80; 17 %) nicht vertraut. Ein kleiner Teil der Teilnehmer gab an, die Daten nur über einen kurzen Zeitraum aufzubewahren.

Bemerkenswert ist, dass diese Frage die höchste Anzahl an Kommentaren (103) erreicht hat. Einige interessante Ergebnisse sollen an dieser Stelle genannt werden:

- Einige Teilnehmer fühlten sich nicht angesprochen und bezeichneten die Grundsätze als nicht anwendbar in ihrem Fachbereich.
- Die Daten werden teilweise auch ohne Kenntnis der Grundsätze aufbewahrt.
- Zur Umsetzung Guter wissenschaftlicher Praxis wird technische und organisatorische Unterstützung benötigt. Die entscheidende Rolle einer Wissenschaftseinrichtung und deren Verantwortung für die fristgemäße Aufbewahrung von Forschungsdaten wird bei kurzen Arbeitsverträgen und Projektlaufzeiten, Jobwechsel oder Ausscheiden in den Ruhestand deutlich sichtbar.
- Manche Datenbestände werden durch Angebote der Fachgemeinschaft archiviert und zugänglich gemacht.
- Die Aufbewahrung bzw. Zugänglichmachung mancher Daten ist durch Lizenz-, Nutzungsverträge, Datenschutz, Urheberrecht oder sonstige Auflagen eingeschränkt.
- Einige Teilnehmer haben sich ausdrücklich dazu bereit erklärt, eine den Grundsätzen entsprechende Infrastruktur zu nutzen.
- Einige Teilnehmer äußerten ihre Bedenken bzgl. der angemessenen Aufbereitung bzw. Dokumentation der Daten, damit deren Entstehung und Interpretation für andere Personen nachvollziehbar gemacht werden können.

Diese Erkenntnisse zeigen, dass der Umgang mit Forschungsdaten ganzheitlich betrachtet werden muss. Damit die Regeln Guter wissenschaftlicher Praxis erfüllt werden können, müssen die Wissenschaftseinrichtungen eine entsprechende Unterstützung anbieten.

3.3.4 Weitere Richtlinien

Anschließend wurden die Teilnehmer gefragt, ob sie weitere formale Richtlinien oder Prozeduren kennen, die den Umgang mit Forschungsdaten genauer bestimmen. Die Ergebnisse werden in der Abbildung 15 dargestellt:

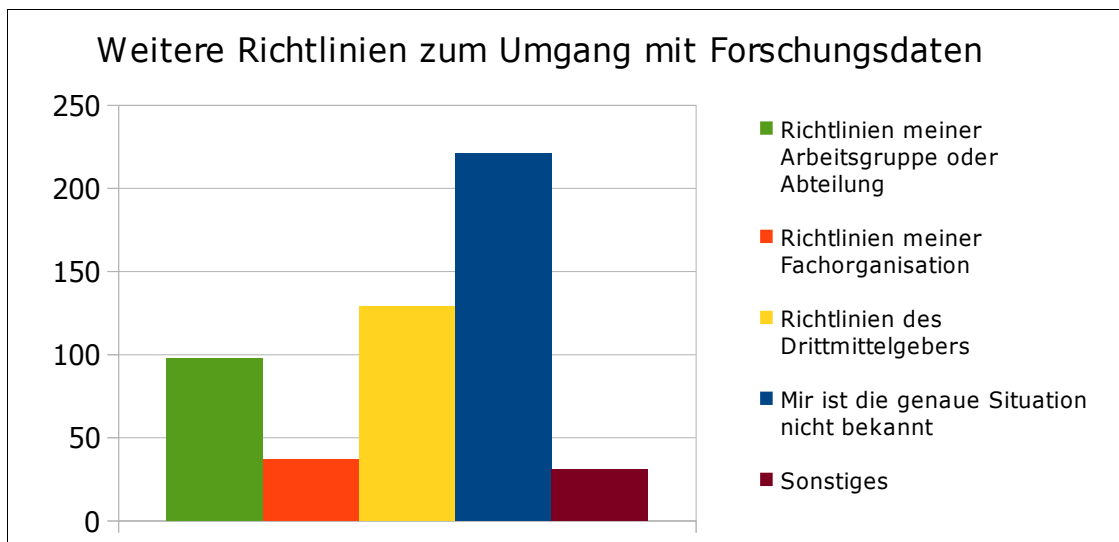


Abbildung 15: Weitere Richtlinien zum Umgang mit Forschungsdaten

Den meisten Teilnehmern war die genaue Situation nicht bekannt (221), weitere bezogen sich auf die Richtlinien des Drittmittelgebers (129) und der eigenen Arbeitsgruppe oder Abteilung (98). Einige gaben auch die Richtlinien der Fachorganisation an (37). In der Kategorie 'Sonstiges' wurden außerdem Datenschutzregelungen, Richtlinien des Datengebers, Vereinbarungen zwischen den Beteiligten, interne Regeln oder ein Ethikvotum genannt. Bei der Entwicklung institutioneller Handlungsempfehlungen zum Umgang mit Forschungsdaten empfiehlt sich die Berücksichtigung der genannten Richtlinien, um direkte Zusammenhänge herzustellen und mögliche Konflikte oder gar Widersprüche frühzeitig zu erkennen.

3.4 Infrastruktur und Service

In den folgenden Unterkapiteln werden die Ergebnisse der Fragengruppe „Infrastruktur und Service“ zur aktuellen und gewünschten Infrastruktur und Unterstützung vorgestellt.

3.4.1 Bereitschaft zur Zugänglichmachung

Zunächst wurden die Teilnehmer gefragt, ob sie grundsätzlich bereit wären, bestimmte Forschungsdaten in einem Datenarchiv abzulegen bzw. zugänglich zu machen. Die Ergebnisse werden in der Abbildung 16 dargestellt:

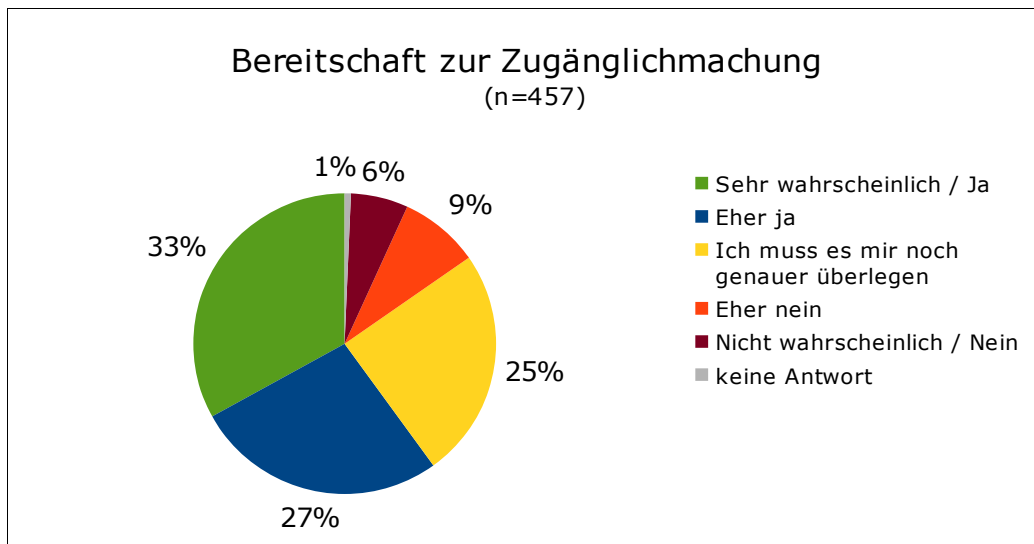


Abbildung 16: Bereitschaft zur Zugänglichmachung

Die Mehrheit der Teilnehmer hat die Frage mit „Sehr wahrscheinlich / Ja“ (151; 33 %) und „Eher ja“ (123; 27 %) positiv beantwortet. Ein weiterer großer Teil der Teilnehmer will es sich noch genauer überlegen (113; 25 %) und nur ein geringer Teil hat sich mit „Eher nein“ (39; 9 %) und „Nicht wahrscheinlich / Nein“ (28; 6 %) negativ geäußert.

Bemerkenswert war wiederum eine hohe Anzahl an Kommentaren (97), die sich in folgende Hinweise bzw. Bedenken gruppieren lassen:

- Bedenken bzgl. des Schutzes der personenbezogenen Daten (u.a. kann eine vorherige Zustimmung der Probanden notwendig sein) und Zugangskontrolle,
- Zugänglichmachung von Forschungsdaten urheberrechtlich oder zum Schutz der Rechte Dritter eingeschränkt oder untersagt,
- Aufwand für geeignete Aufbereitung bzw. Zweifel an Nachnutzung bzw. Nutzen für andere,
- Begrüßung des Open-Access-Gedankens und weiteres Interesse daran,
- Forschungsdaten werden bevorzugt erst nach der eigenen Auswertung bzw. Publikation der Forschungsergebnisse anderen zugänglich gemacht,
- Manche Teilnehmer gaben an, sie produzieren keine Forschungsdaten bzw. konnten die Frage für ihren Fachbereich nicht anwenden,
- Zugänglichmachung von Forschungsdaten soll in einem benutzerfreundlichen Datenarchiv erfolgen bzw. die Arbeitsabläufe unterstützen,
- zur Freigabe ist die Absprache mit dem Leiter oder Betreuer notwendig,
- manche Teilnehmer sehen Forschungsdaten als ihr Eigentum bzw. stehen unter zu hohem Konkurrenzdruck, um diese vorzeitig öffentlich zugänglich zu machen.

Bemerkenswert ist, dass sich diese Frage bewusst auf „bestimmte Forschungsdaten“ bezog. Dem lag die Annahme der Autoren zugrunde, dass nicht alle Forschungsdaten öffentlich zugänglich gemacht werden dürfen oder sollen. Einige damit verbundene Aspekte konnten mittels der folgenden zwei Fragen detaillierter betrachtet werden.

3.4.2 Forschungsdaten in einem Datenarchiv

Im Zusammenhang mit der Bereitschaft zur Zugänglichmachung von Forschungsdaten wurden die

Teilnehmer gefragt, welche Forschungsdaten sie in einem Datenarchiv ablegen bzw. zugänglich machen würden. Die Ergebnisse werden in der Abbildung 17 dargestellt:

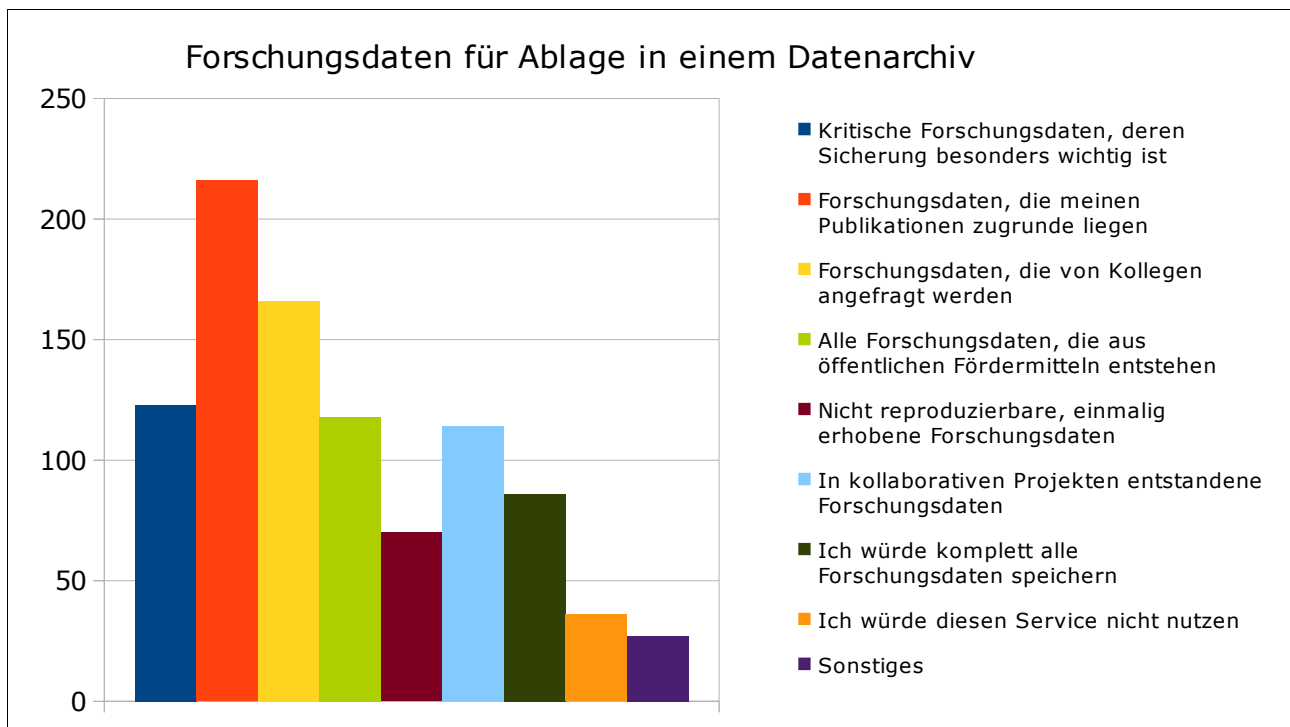


Abbildung 17: Forschungsdaten für Ablage in einem Datenarchiv

Die Teilnehmer würden größtenteils die Forschungsdaten ablegen, die ihren Publikationen zugrunde liegen (216) und die von Kollegen angefragt werden (166). Zu fast gleichen Anteilen wurden kritische Forschungsdaten, deren Sicherung besonders wichtig ist (123), alle Forschungsdaten, die aus öffentlichen Mitteln entstehen (118) sowie in kollaborativen Projekten entstandene Forschungsdaten (114) genannt. Erwähnenswert ist auch die Häufigkeit der beiden gegensätzlichen Positionen: Während einige Teilnehmer alle Forschungsdaten speichern würden (86), würden andere diese Möglichkeit gar nicht wahrnehmen (36). Unter 'Sonstiges' bezogen sich die meisten Teilnehmer auf die Kommentare zur Frage davor und gaben nachnutzbare, nicht vertrauliche Forschungsdaten und die Kontrolle über den Zugriff darauf als Gründe an.

Zusammenfassend lässt sich sagen, dass die Teilnehmer ihre Forschungsdaten vor allem zum Zweck der Unterstützung, des Austausches und der Verbreitung der Forschungsergebnisse zugänglich machen würden. Bei genauerer Analyse der potentiellen „Nicht-Nutzer“ hat es sich zudem herausgestellt, dass es sich dabei vor allem um Teilnehmer handelt, die ihre Arbeitsmaterialien selbst nur schwerlich als Forschungsdaten bezeichnen würden (z.B. Theoretische Fachbereiche) oder deren Arbeitsmethoden strengen Vorschriften oder Beschränkungen unterliegen (z.B. Datenschutz).

3.4.3 Bevorzugtes Datenarchiv

Ferner wurden die Teilnehmer gefragt, welches Datenarchiv sie für die Ablage der Forschungsdaten bzw. deren Zugänglichmachung am ehesten nutzen würden. Die Ergebnisse werden in der Abbildung 18 dargestellt:

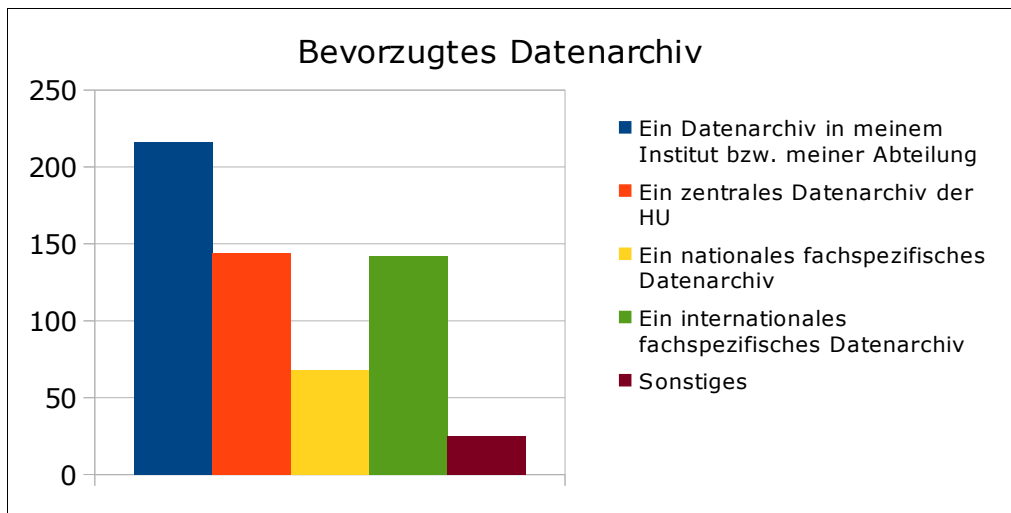


Abbildung 18: Bevorzugtes Datenarchiv

Die höchste Zahl der Antworten hat ein Datenarchiv in eigenem Institut bzw. in der eigenen Abteilung erreicht (216), gefolgt von einem zentralen Datenarchiv der HU (144) und einem fachspezifischen internationalen Datenarchiv (142). Eher selten hingegen wird ein nationales fachspezifisches Datenarchiv bevorzugt (68). Die Nutzung solcher Datenarchive kann dabei mit dem Zweck des Forschungsvorhabens (z.B. für landesweite Studien) verbunden sein. Die genaue Rolle von nationalen fachspezifischen Datenarchiven wurde jedoch im Rahmen dieser Umfrage nicht näher untersucht.

Grundsätzlich haben sich während der Umfrage drei unterschiedliche Szenarien für die Nutzung von Datenarchiven oder Repositorien herauskristallisiert:

- Operative Speicherung von Forschungsdaten: Die Forschungsdaten sollen als aktuelle Arbeitsmaterialien laufend gespeichert und mit den Kooperationspartnern ausgetauscht und gemeinsam bearbeitet werden können. Dies würde beispielsweise den Austausch der Daten als E-Mail-Anhang oder per Post ersparen. Entscheidend dabei ist, dass der Zugriff auf bestimmte Personen beschränkt wird. Für dieses Szenario werden tendenziell interne Datenarchive der eigenen Einrichtung bzw. der Einrichtung des Partners bevorzugt;
- Langzeitarchivierung von Forschungsdaten: Im Sinne der Guten wissenschaftlichen Praxis sollten die Forschungsdaten nach Abschluss eines Projektes oder einer Forschungstätigkeit in einem (institutionellen) Repository oder Datenarchiv aufbewahrt werden. Dies kann unabhängig oder im Zusammenhang mit der Veröffentlichung von Forschungsdaten erfolgen;
- Veröffentlichung von Forschungsdaten: Forschungsdaten sollen zur Unterstützung der publizierten Forschungsergebnisse (z.B. eines Zeitschriftenartikels) veröffentlicht werden. In diesem Fall könnten die Forschungsdaten mit weiteren Objekten verknüpft und kontextualisiert werden. Entscheidend ist dabei die Auffindbarkeit und die Nachvollziehbarkeit der publizierten Forschungsdaten, die beispielsweise durch eine ausführliche Dokumentation bzw. Metadaten und die Unterbringung der Forschungsdaten in den der Fachöffentlichkeit bekannten Datenarchiven oder Repositorien gewährleistet werden kann. Sollten die Forschungsdaten als eigenständige Objekte publiziert werden, nimmt die Bedeutung einer klaren Kennzeichnung der Autorschaft und die Zitierbarkeit der Forschungsdaten zu. Da die Forschung zunehmend durch weltweite Kooperationen geprägt ist, werden für dieses Szenario tendenziell fachspezifische internationale Datenarchive bevorzugt.

3.4.4 Hilfreiche Kriterien

Des Weiteren wurden die Teilnehmer befragt, nach welchen Kriterien sie ein für ihre Forschungsdaten geeignetes Datenarchiv auswählen würden. Die Ergebnisse werden in den Abbildungen 19a - e dargestellt:

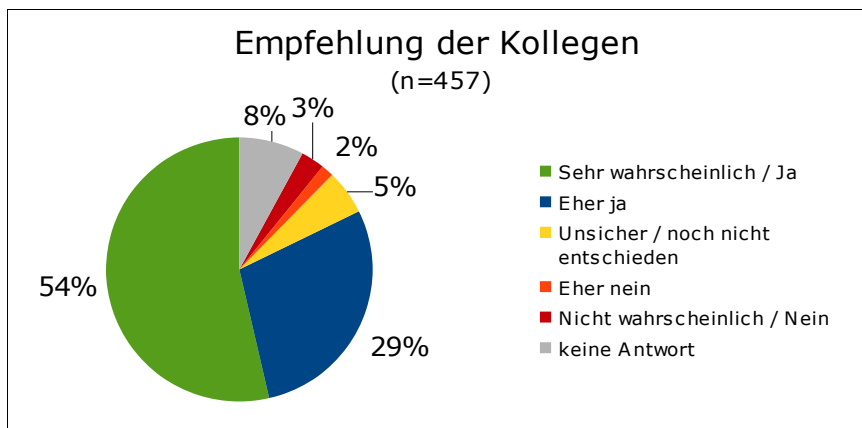


Abbildung 19a: Hilfreiche Kriterien: Empfehlung der Kollegen

Die Ergebnisse der ersten Teilfrage zeigen, dass der Empfehlung der Kollegen ein sehr hoher Wert zugemessen wird: Die meisten Teilnehmer bewerteten die Orientierung daran als sehr wahrscheinlich (245; 54 %) oder eher wahrscheinlich (131; 29 %). Einige Teilnehmer waren unsicher oder noch nicht entschieden (25; 5 %) und nur wenige bezeichneten das Kriterium als eher unwahrscheinlich (7; 2 %) oder nicht wahrscheinlich (13; 3 %) für ihre Entscheidung.

Dieses Ergebnis zeigt zudem, dass bei der Umsetzung eines institutionellen Konzeptes zum Umgang mit Forschungsdaten die Gewinnung von sogenannten „local champions“ besonders vielversprechend sein könnte.⁴⁴ Diese Gruppe von engagierten Fachwissenschaftlern an einzelnen Einrichtungen soll daher schon bei der Entwicklung der unterstützenden Serviceangebote möglichst früh einbezogen werden, um diese bedürfnisorientiert zu gestalten und dadurch die Akzeptanz zu steigern.

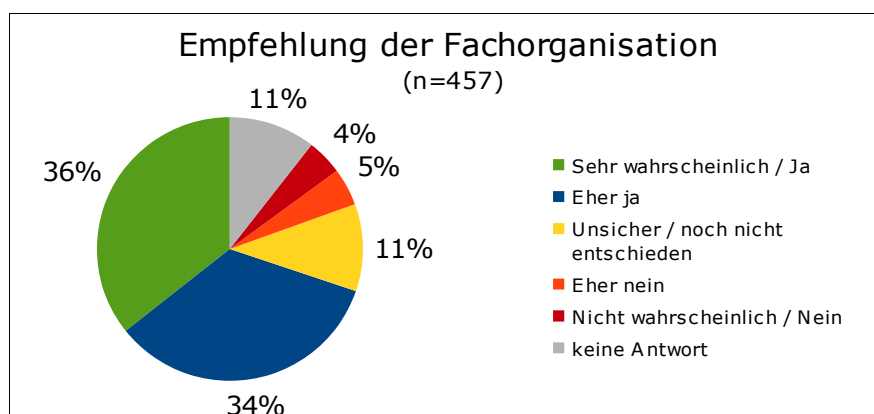


Abbildung 19b: Hilfreiche Kriterien: Empfehlung der Fachorganisation

Ähnlich hoch eingeschätzt wird die Empfehlung der Fachorganisation: Die meisten Teilnehmer würden sich sehr wahrscheinlich (163; 36 %) oder eher wahrscheinlich (156; 34 %) daran orientieren. Ein Teil der Teilnehmer war noch nicht entschieden (49; 11 %), andere bewerteten dieses Kriterium als eher unwahrscheinlich (21; 5 %) oder nicht wahrscheinlich (20; 4 %). Das Ergebnis deutet darauf hin, dass die Fachorganisationen zu einer treibenden Kraft bei der Etablierung fachgerechter Praktiken sind, um Forschungsdaten zugänglich zu machen.

44 Vgl. Freiman et al. (2010)

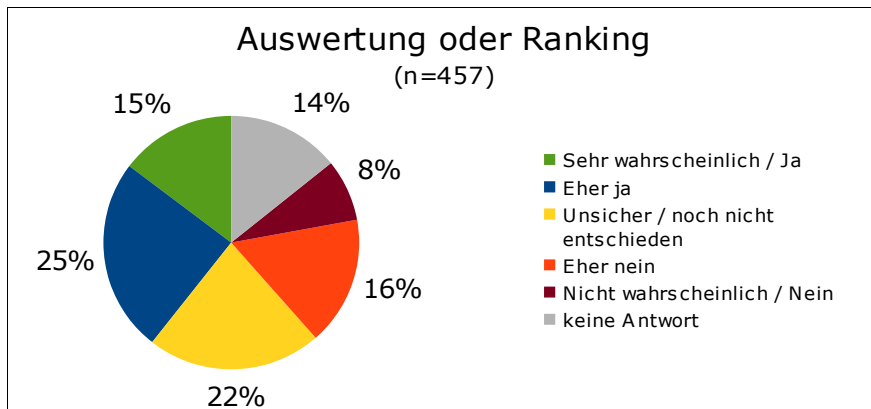


Abbildung 19c: Hilfreiche Kriterien: Auswertung oder Ranking

Bezüglich der Bedeutung von Auswertungen oder Rankings in unabhängigen Verzeichnissen bei der Auswahl eines Datenarchivs, waren die Meinungen der Teilnehmer etwa gleich verteilt. Als sehr wahrscheinlich relevant für die Entscheidungsfindung bewerteten dieses Kriterium 67 (15 %) der Teilnehmer, als eher wahrscheinlich 113 (25 %). Noch unsicher waren 101 (22 %) Teilnehmer, als eher unwahrscheinlich schätzten 75 (16 %) bzw. als nicht wahrscheinlich 36 (8 %) der Teilnehmer dieses Kriterium für die Entscheidungsfindung ein.

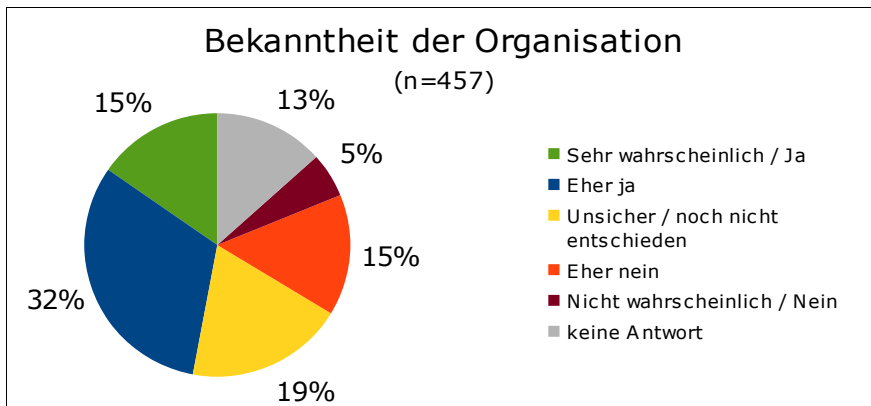


Abbildung 19d: Hilfreiche Kriterien: Bekanntheit der Organisation

Weiterhin wurden die Teilnehmer gefragt, ob die Bekanntheit der Organisation, die das Datenarchiv betreibt, eine Bedeutung bei dessen Auswahl spielen würde. Die Einschätzungen der Teilnehmer fielen ähnlich wie bei der Frage zuvor aus. Als sehr wahrscheinlich bzw. eher wahrscheinlich bewerteten es 70 (15 %) bzw. 145 (32 %) der Teilnehmer. Weitere 88 (19 %) waren unsicher oder noch nicht entschieden, 68 (15 %) antworteten mit „eher unwahrscheinlich“ und 25 (5 %) mit „nicht wahrscheinlich“.

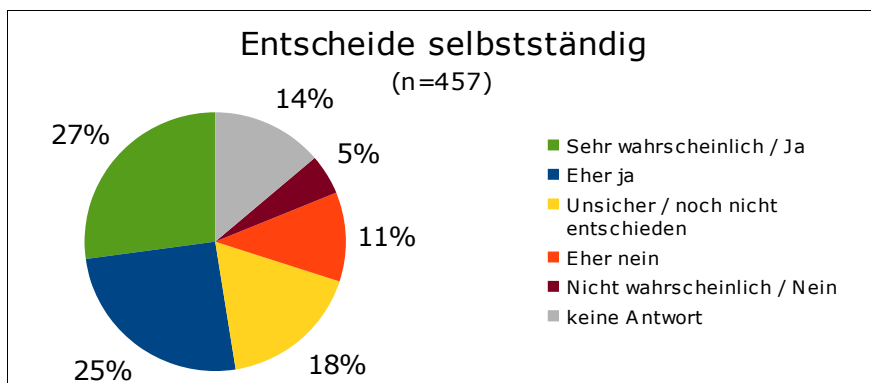


Abbildung 19e: Hilfreiche Kriterien: Entscheide selbstständig

Letztlich sollte geklärt werden, ob die Teilnehmer ohne Hilfestellungen auskommen und über die Auswahl eines Datenarchivs am ehesten selbstständig entscheiden würden. Etwa die Hälfte der Teilnehmer bewertete ein solches Vorgehen als sehr wahrscheinlich (124; 27 %) oder eher wahrscheinlich (116; 25 %). Darüber hinaus schätzten dies die Teilnehmer als „unsicher“ oder „noch nicht entschieden“ (80; 18 %) und „eher unwahrscheinlich“ (51; 11 %) und „nicht wahrscheinlich“ (23; 5 %) ein.

In der Abbildung 20 werden die Ergebnisse aller fünf Teilfragen in einer Matrix visualisiert:

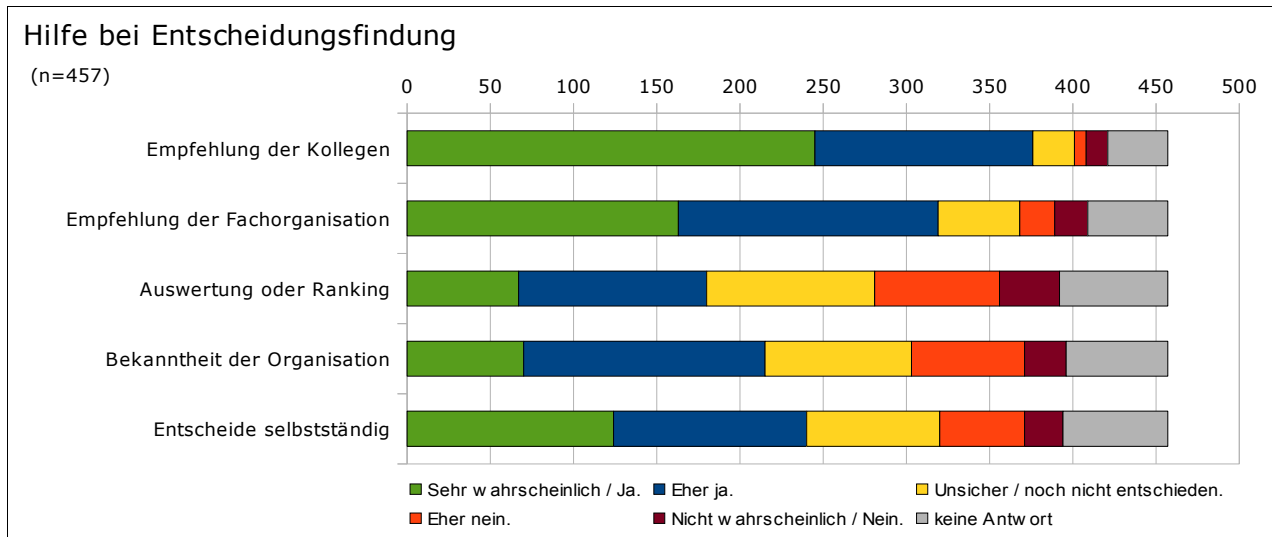


Abbildung 20: Hilfe bei Entscheidungsfindung (Matrix)

Interessanterweise deuten die Antworten auf die letzte Teilfrage darauf hin, dass die selbstständige Entscheidung nicht der Empfehlung der Kollegen oder der Empfehlung der Fachorganisation vorgezogen wird. Es würden sich mehr Teilnehmer auf diese Empfehlungen verlassen, als eine Entscheidung ohne Bezugnahme auf weitere Aspekte zu treffen. Dies kann der Tatsache geschuldet sein, dass für die meisten Teilnehmer die Ablage von Forschungsdaten in Datenarchiven bzw. deren Veröffentlichung ein neues Aufgabengebiet ist (s. auch 3.3.2 Veröffentlichung und Nachnutzung), was eine umfangreiche Beratung und abgestimmte Ansatzpunkte der inhaltlichen Auseinandersetzung erfordert. Die Unsicherheit wird möglicherweise auch dadurch signalisiert, dass ein im Vergleich recht hoher Anteil von Teilnehmern (8-14 %) die Kriterien in Teilfragen nicht bewertet hat (keine Antwort).

Gleichzeitig spielen Auswertungen oder Rankings und die Bekanntheit der betreibenden Organisation den Ergebnissen zufolge eine geringere Rolle bei der Auswahl eines Datenarchivs. Das kann als ein Indiz dafür gewertet werden, dass eine gewisse Vertrauensbasis notwendig ist. Allein die Bekanntheit der Organisation oder das Ranking nach bestimmten Kriterien ist möglicherweise nicht ausreichend aussagekräftig. Die Beratung mit Kollegen oder durch die Fachorganisation hingegen bedeutet oft eine persönliche Kontaktstelle, mit der die Forschenden bereits gut vertraut sind.

Zusätzlich bot sich die Möglichkeit, sonstige relevante Kriterien zu benennen. Dabei wurden u.a. die Unterstützung des kollaborativen Arbeitens (insbesondere durch den Zugang für externe Partnerorganisationen) und fachspezifischer Arbeitsabläufe, klare Nutzungsrichtlinien, ein nicht-kommerzieller Hintergrund, die Möglichkeit der Langzeitarchivierung und der Zugangskontrolle sowie eine Beurteilung durch IT-Experten genannt.

3.4.5 Nutzung anderweitiger Infrastruktur

Des Weiteren wurden die Teilnehmer gefragt, ob sie die technische Infrastruktur anderer Einrichtungen nutzen (z.B. in gemeinsamen Projekten). Die Ergebnisse werden in der Abbildung 21 dargestellt:

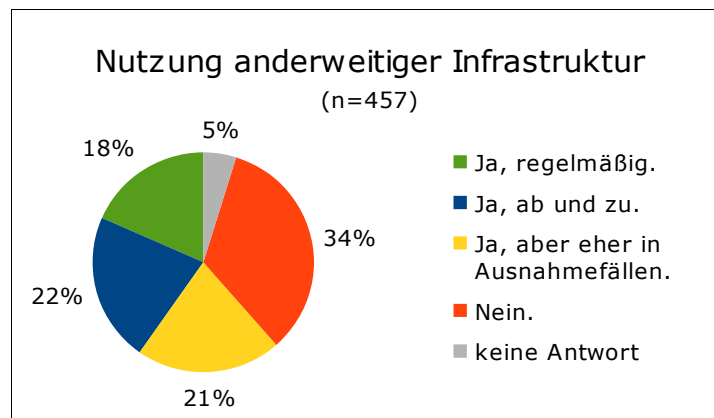


Abbildung 21: Nutzung anderweitiger Infrastruktur

Die Antworten der Teilnehmer verteilten sich wie folgt: Ja, regelmäßig (84; 18 %), Ja, ab und zu (100; 22 %), Ja, aber eher in Ausnahmefällen (97; 21 %), Nein (154; 34 %) und keine Antwort (22; 5 %). Die Teilnehmer benannten zudem konkrete fachspezifische Angebote von Kooperationspartnern (u.a. im Rahmen eines Sonderforschungsbereichs, eines Exzellenzclusters, in Drittmittelprojekten und von außeruniversitären Wissenschaftsorganisationen sowie Cloud-Dienste. Die Frage diente dazu, die von den Forschenden der Humboldt-Universität benutzte Infrastruktur teilweise zu identifizieren, damit diese Erkenntnisse bei der Entwicklung von eigenen Angeboten berücksichtigt werden können.

3.4.6 Gewünschte Unterstützung

Schließlich wurden die Teilnehmer gefragt, welche Serviceleistungen sie sich von der Humboldt-Universität wünschen würden. Die Ergebnisse werden in der Abbildung 22 dargestellt:

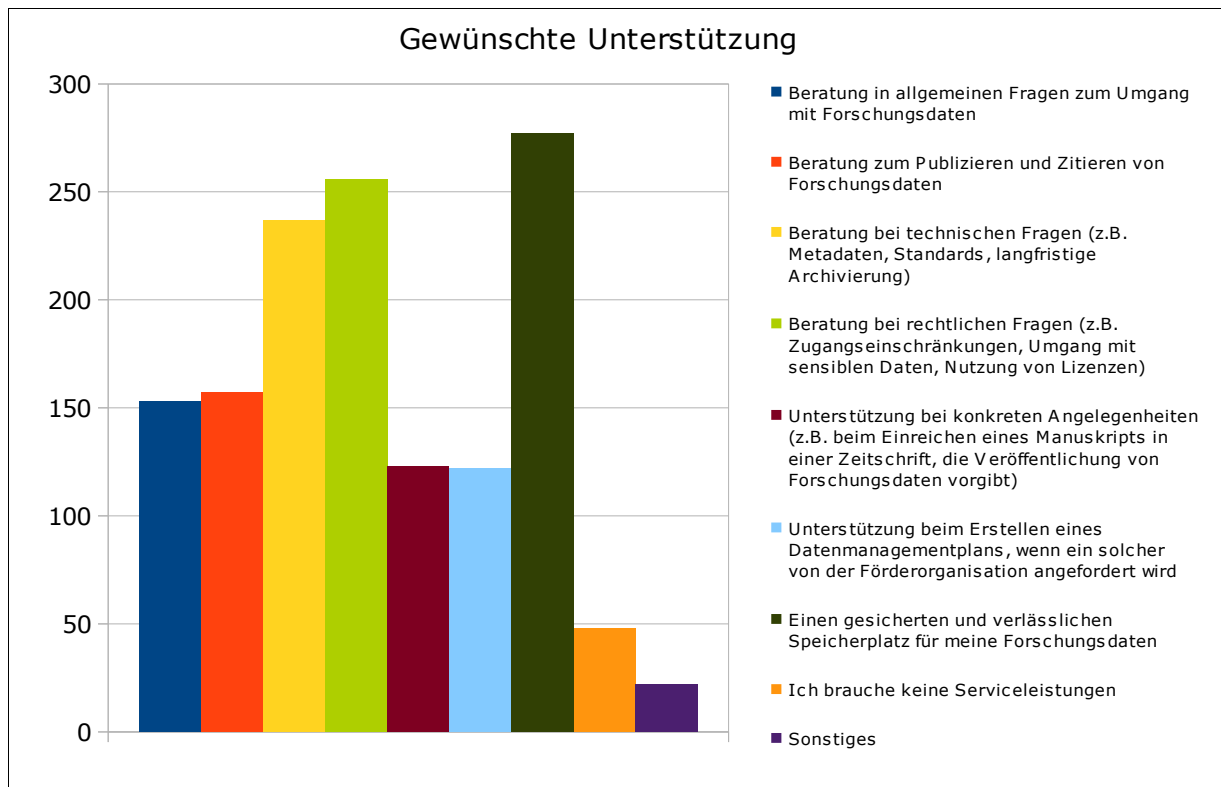


Abbildung 22: Gewünschte Unterstützung

Die drei Spitzenpositionen nahmen dabei mit großem Abstand folgende Antworten ein:

- Einen gesicherten und verlässlichen Speicherplatz für meine Forschungsdaten (277),
- Beratung bei rechtlichen Fragen (z.B. Zugangseinschränkungen, Umgang mit sensiblen Daten, Nutzung von Lizenzen) (256) und
- Beratung bei technischen Fragen (z.B. Metadaten, Standards, langfristige Archivierung) (237).

Die Liste setzt sich fort mit:

- Beratung zum Publizieren und Zitieren von Forschungsdaten (157),
- Beratung in allgemeinen Fragen zum Umgang mit Forschungsdaten (153),
- Unterstützung bei konkreten Angelegenheiten (z.B. beim Einreichen eines Manuskripts in einer Zeitschrift, die Veröffentlichung von Forschungsdaten vorgibt) (123),
- Unterstützung beim Erstellen eines Datenmanagementplans, wenn ein solcher von der Förderorganisation angefordert wird (122).

Unter 'Sonstiges' wurden u. a. die Unterstützung in individuellen Belangen, ein direkter Ansprechpartner, bestimmte Hardware und ein eigener Cloud-Dienst genannt. 48 Teilnehmer gaben an, gar keine Serviceleistungen zu benötigen. Bei genauerer Analyse waren dies diejenigen Fachbereiche, die ihre Arbeitsweise nicht mit dem Begriff „Forschungsdaten“ verbinden konnten (begrifflich-theoretische oder interpretierende Forschungsformen) und diejenigen, die durch bestehende Infrastruktur bereits gut versorgt sind (z.B. Forscher, die in großen Kooperationsprojekten arbeiten). Gleichzeitig wurde durch einige Teilnehmer dennoch angemerkt, dass eine rechtliche Beratung und ein direkter Ansprechpartner, der bei Bedarf aufgesucht werden könnte, für sie hilfreich sein könnte.

Zusammenfassend kann gesagt werden, dass vor allem eine pragmatische Unterstützung bei der Beantwortung grundlegender technischer, organisatorischer und rechtlicher Fragen benötigt wird. Weiterreichende Angelegenheiten wie das Zitieren von Forschungsdaten oder die Erstellung eines von vielen Förderorganisationen in anderen Ländern vorgeschriebenen Datenmanagementplans können sinnvollerweise erst dann vorangebracht werden, wenn diese sich auf feste Grundlagen beziehen können. Insbesondere wird aktuell eine rechtliche Aufklärung gefragt, die die Bündelung informationswissenschaftlicher und rechtswissenschaftlicher Expertise voraussetzt. Aufgrund der Vielschichtigkeit des Themas wird außerdem die Etablierung einer zentralen Anlaufstelle als vorteilhaft gesehen.

3.5 Weiteres Interesse

Auch wurde den Teilnehmern die Möglichkeit gegeben, weiteres Interesse zu bekunden. Dies betraf die Ergebnisse der Umfrage sowie die Teilnahme an einem individuellen Interview. Dabei meldeten sich 82 Teilnehmer für ein Interview, bei dem sie von ihren Erfahrungen zum Umgang mit Forschungsdaten und Erwartungen an zentrale Serviceleistungen berichten könnten. 129 Teilnehmer waren an den Ergebnissen der Umfrage interessiert. Derzeit findet die Durchführung von Interviews statt.⁴⁵

45 Für die Methodik zur Durchführung von Zielgruppenanalysen s.z.B. Quirk et al. (2008)

4 Vergleich der Antworten zwischen Teilnehmergruppen

Die Ergebnisse der Umfrage werden im folgenden Kapitel zwischen den zwei größten Teilnehmergruppen „Professor(in)“ und „Wissenschaftliche(r) Mitarbeiter(in)“ verglichen. Diesem Teil des Berichts liegt die Auswertung der Ergebnisse in dem Tabellendokument „Umfrage_Forschungsdaten_2013-Vergleich_nach_Groupen.ods“ zugrunde.

Dabei sollen nicht mehr die Antworten zu allen Fragen präsentiert werden, sondern nur diejenigen, bei denen sich Unterschiede zwischen den beiden Gruppen ergeben haben. Das Ziel des Vergleichs soll nicht sein, „Klischees zu bedienen“, sondern vielmehr die Tendenzen aufzuzeigen.

Der Prozentanteil wird dabei nach der Gesamtzahl der Teilnehmer in der jeweiligen Gruppe berechnet (s. Tab. 2):

Teilnehmergruppe	Anzahl der Teilnehmer
"Professor(in)"	123
"wissenschaftliche(r) Mitarbeiter(in)"	267

Tabelle 2: Teilnehmerzahl nach Gruppen

4.1 Datentypen

Die Antworten der Fragen zu den Quellen von Forschungsdaten, deren Typen, Speicherorten, zusätzlichen Datenträgern unterschieden sich nur unwesentlich. Bei der Gesamtgröße der Forschungsdaten gab es jedoch einige Unterschiede. Die Ergebnisse werden in der Abbildung 23 dargestellt:

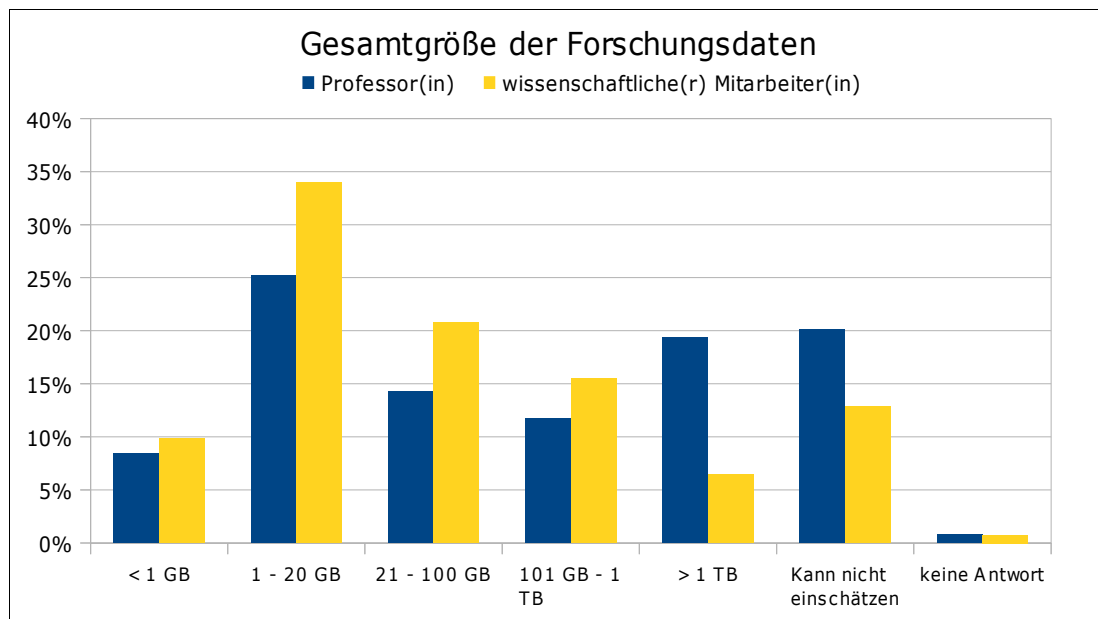


Abbildung 23: Gesamtgröße der Forschungsdaten (Vergleich)

Wie aus der Abbildung 23 hervorgeht, gaben wissenschaftliche Mitarbeiter die Gesamtgröße ihrer Forschungsdaten öfter mit niedrigeren Werten an als Professoren. Dies ist insbesondere bei den Antworten „1 - 20 GB“ sowie „> 1 TB“ sichtbar. Gleichzeitig konnten mehr Professoren die Gesamtgröße nicht einschätzen. Das Ergebnis kann möglicherweise dadurch erklärt werden, dass mit der längeren Zeit in der Forschung auch die Menge an Forschungsdaten größer wird. Zudem mag es sein, dass die Pflege von Forschungsdaten an weitere Personen delegiert wird bzw. deren Gesamtgröße nicht mehr übersichtlich ist.

In der Abbildung 24 wird der Vergleich bei der Erstellung der Sicherungskopien dargestellt:

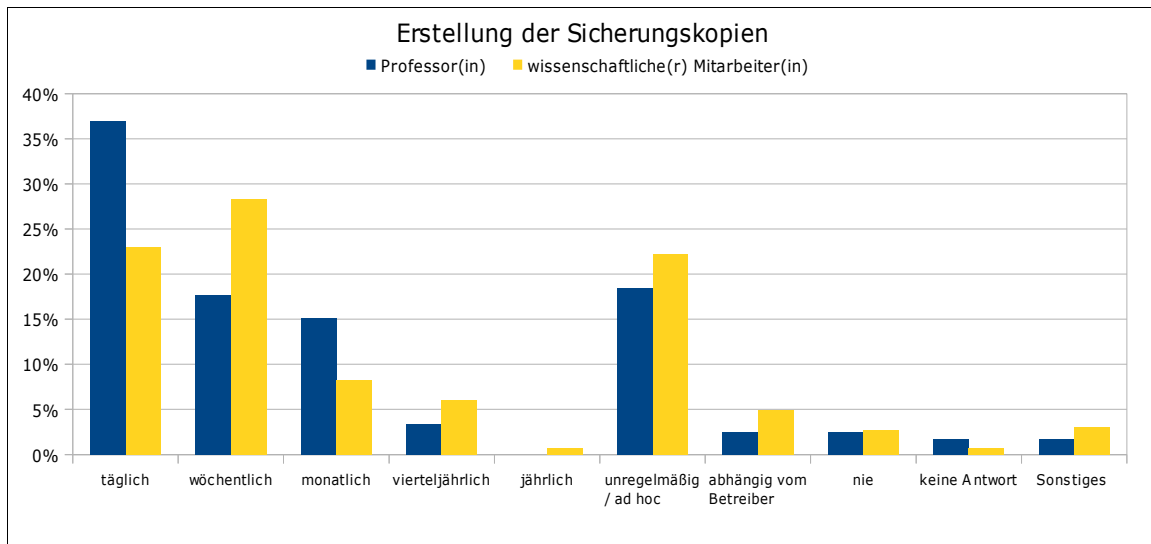


Abbildung 24: Erstellung der Sicherungskopien (Vergleich)

Wie aus dem Vergleich hervorgeht, werden die Forschungsdaten von Professoren öfter täglich oder monatlich gesichert. Dazu gaben wissenschaftliche Mitarbeiter tendenziell öfter an, ihre Forschungsdaten wöchentlich oder unregelmäßig zu sichern. Bei dieser Frage wurde jedoch nicht geklärt, ob alle oder nur bestimmte Forschungsdaten (z.B. neulich bearbeitete) gesichert werden.

In der Abbildung 25 wird die Verantwortung für Forschungsdaten verglichen:

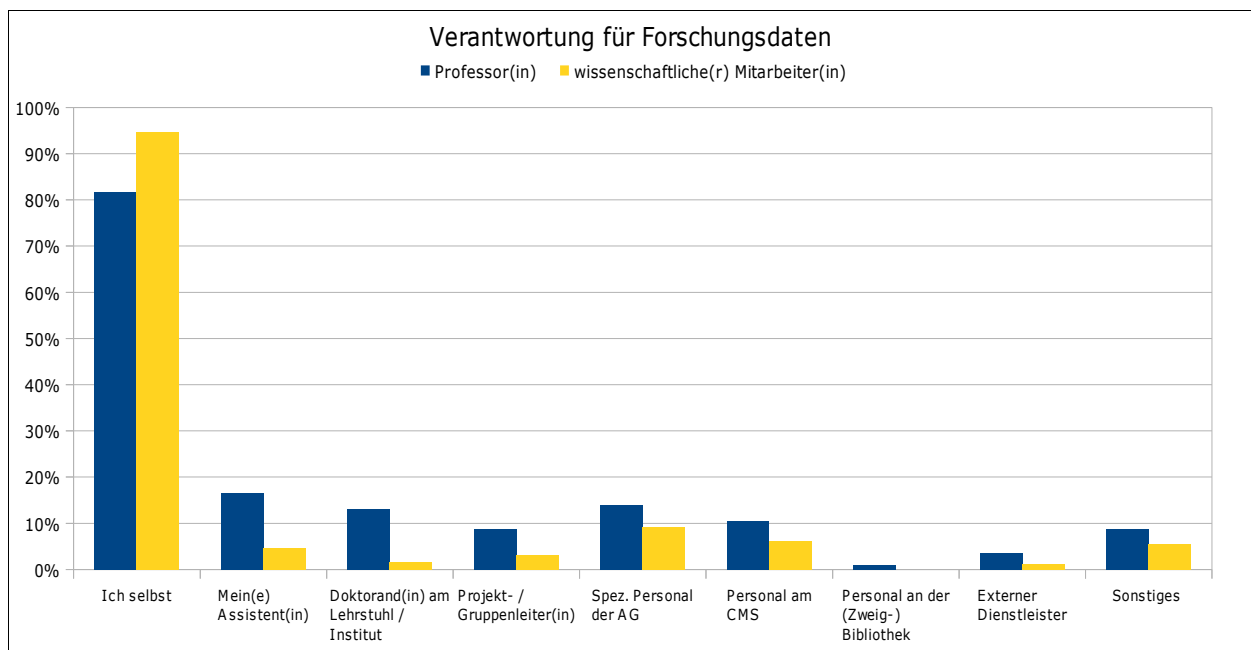


Abbildung 25: Verantwortung für Forschungsdaten (Vergleich)

Auch wenn beide Teilnehmergruppen überwiegend selbst für ihre Forschungsdaten verantwortlich sind, nehmen die Professoren den Ergebnissen zufolge die Unterstützung durch andere Personen öfter in Anspruch. Außerdem liegt es nahe, dass die Professoren häufiger über Assistenten verfügen, an die sie diese Aufgabe delegieren könnten.

4.2 Zugang

In dieser Fragensgruppe wird zunächst der Zugriff auf Forschungsdaten verglichen. Die Ergebnisse werden in der Abbildung 26 dargestellt:

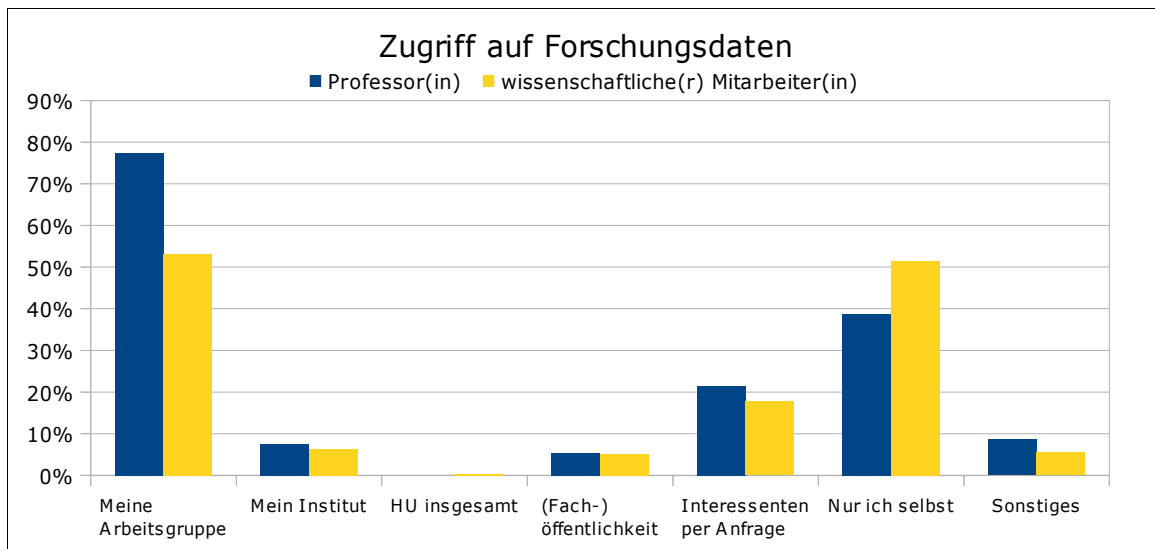


Abbildung 26: Zugriff auf Forschungsdaten (Vergleich)

Aus den Ergebnissen lässt sich erkennen, dass der Zugriff auf Forschungsdaten der eigenen Arbeitsgruppe von wissenschaftlichen Mitarbeitern seltener gewährt wird als von den Professoren. Gleichzeitig gaben wissenschaftliche Mitarbeiter öfter an, sie haben alleinig Zugriff auf Fortschreibungsdateien⁴⁶. Dies kann teilweise dadurch erklärt werden, dass einige wissenschaftliche Mitarbeiter an eigenen Dissertationsvorhaben arbeiten (s. auch 3.1.2 Teilnehmergruppen) und ihre Forschungsergebnisse noch nicht veröffentlicht haben. Das Ergebnis mag sich jedoch von Fach zu Fach unterscheiden, was aber an dieser Stelle nicht genauer untersucht wurde.

⁴⁶ Bei dieser Frage wurden Mehrfachantworten erlaubt, so dass die Summe der Antworten in beiden Gruppen 100 % übersteigt. Bei einer künftigen Untersuchung wäre es jedoch angebracht, die Antwortoption „Nur ich selbst“ von den anderen zu trennen.

Im Folgenden wird die Erfahrung in Veröffentlichung und Nachnutzung von Forschungsdaten verglichen. Die Ergebnisse werden in Abbildung 27a - c dargestellt:

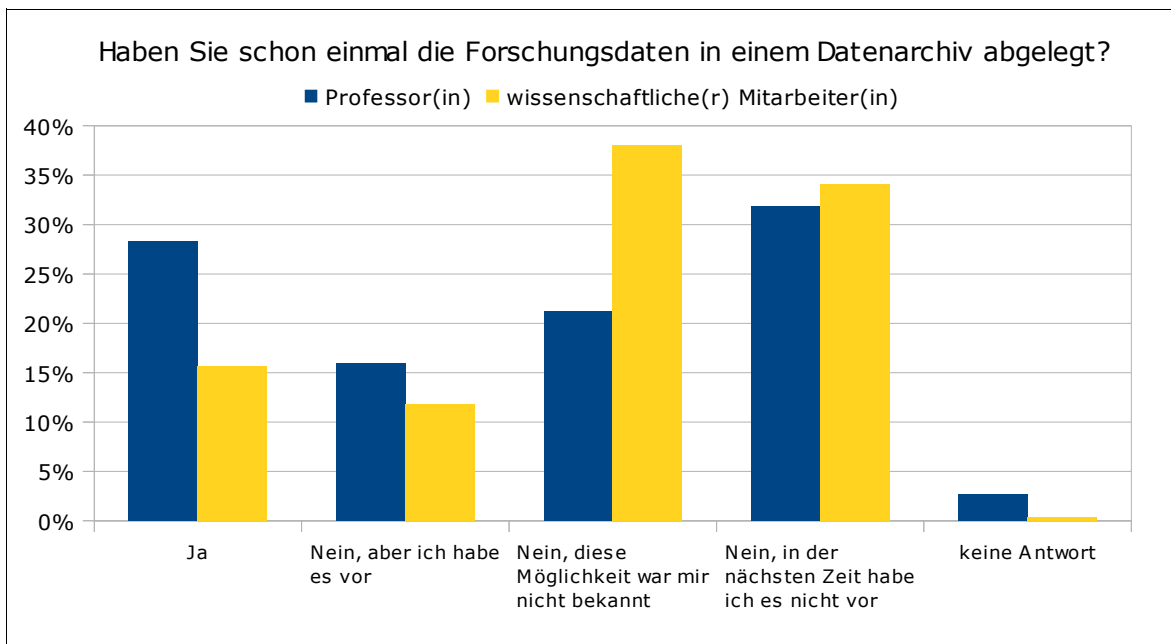


Abbildung 27a: In einem Datenarchiv abgelegt? (Vergleich)

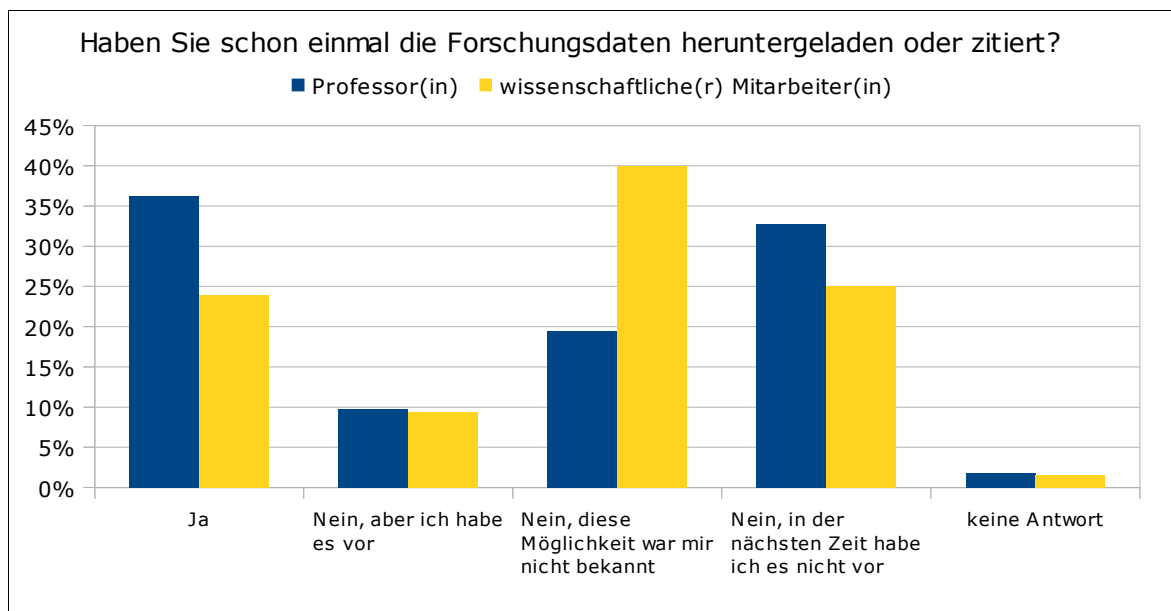


Abbildung 27b: Heruntergeladen oder zitiert? (Vergleich)

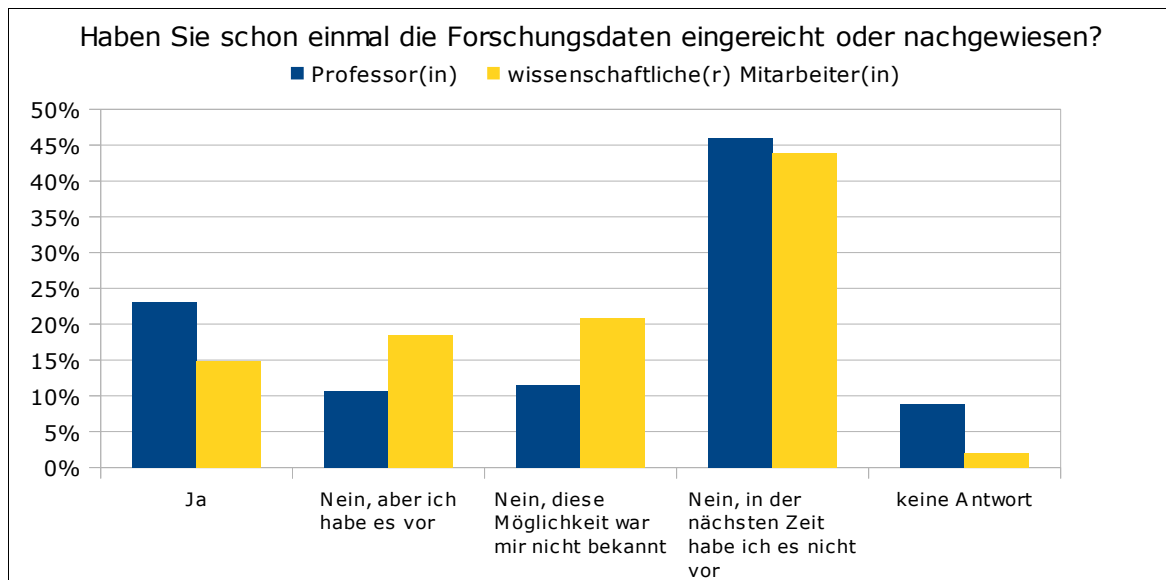


Abbildung 27c: Eingereicht oder nachgewiesen? (Vergleich)

Aus den Ergebnissen geht hervor, dass die Möglichkeiten der Veröffentlichung und Nachnutzung von Forschungsdaten von den Professoren eher genutzt werden und den wissenschaftlichen Mitarbeitern oftmals nicht bekannt sind. Interessanterweise war der Anteil der Teilnehmer, die ihre Forschungsdaten zusammen mit dem Manuskript einzureichen oder nachzuweisen vorhatten, unter den wissenschaftlichen Mitarbeitern größer als bei den Professoren (Abbildung 27c). Dieses Ergebnis kann möglicherweise auf ein größeres Interesse für die Veröffentlichung von Forschungsdaten in dieser Karrierestufe hindeuten, einhergehend mit einem höheren Unterstützungsbedarf beim Publizieren und Zitieren sowie konkreten Angelegenheiten im Umgang mit Forschungsdaten (vgl. Abbildung 34).

Ferner werden in der Abbildung 28 die Antworten zur Sicherung guter wissenschaftlicher Praxis verglichen:

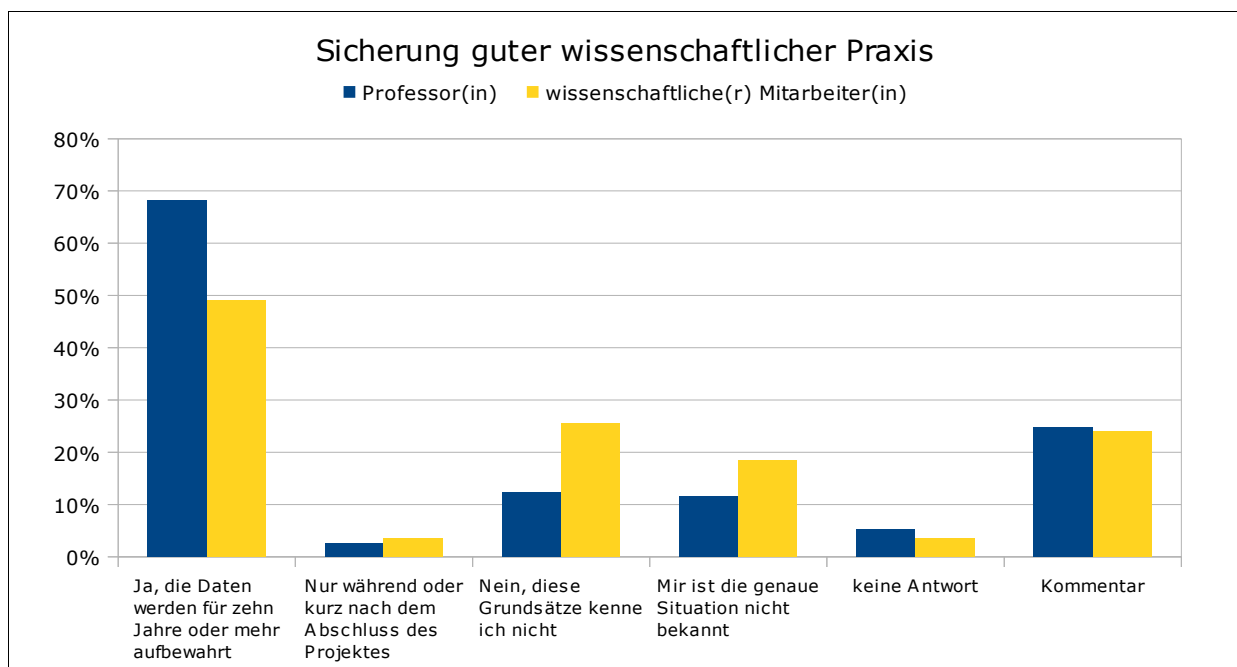


Abbildung 28: Sicherung guter wissenschaftlicher Praxis (Vergleich)

Aus den Ergebnissen wird deutlich, dass diese Grundsätze unter wissenschaftlichen Mitarbeitern öfter unbekannt sind bzw. von ihnen seltener eingehalten werden (können). Ausführliche

Kommentare dazu finden sich in Kapitel 3.3.3 Sicherung guter wissenschaftlicher Praxis.

Daneben sollte geklärt werden, welche weitere Richtlinien zum Umgang mit Forschungsdaten den Teilnehmern bekannt sind. Die Ergebnisse werden in der Abbildung 29 dargestellt:

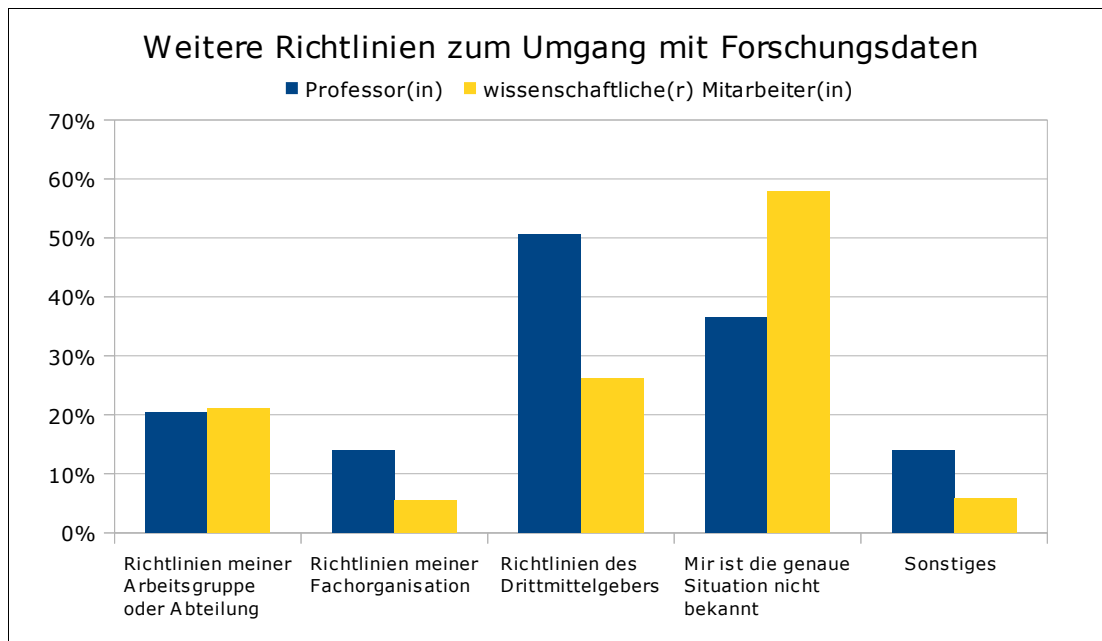


Abbildung 29: Weitere Richtlinien zum Umgang mit Forschungsdaten (Vergleich)

Auch bei dieser Frage zeigte es sich, dass die mit Forschungsdaten verbundenen Regularien den wissenschaftlichen Mitarbeitern öfter unbekannt sind. Möglicherweise hängt dies damit zusammen, dass die Verantwortung für die Einhaltung der Vorgaben durch leitende Personen bzw. Professoren übernommen wird. Allerdings sollten die Erwartungen mancher Drittmittelgeber (z.B. Hinweise für Antragstellung⁴⁷ der DFG) effizienter kommuniziert werden. Viel Potential gibt es auch bei der Weiterentwicklung und Etablierung fachspezifischer Richtlinien (z.B. durch Fachorganisationen).

4.3 Infrastruktur und Service

In dieser Fragengruppe wird zunächst die Bereitschaft zur Zugänglichmachung verglichen. Die Ergebnisse werden in der Abbildung 30 dargestellt:

47 Vgl. Unter „Hinweise für Antragsteller – Hinweise für die Planungsphase“ heißt es: „Werden in dem geplanten Projekt systematisch Forschungsdaten erhoben, die für eine Nachnutzung geeignet sind, sollte der Antrag ein entsprechendes Konzept für die Überführung der Forschungsdaten in vorhandene Datenbanken oder Repositorien enthalten. Häufig ist es in diesem Fall sinnvoll, bereits in der Planungsphase Kontakt zu den Ansprechpersonen der entsprechenden Infrastrukturen aufzunehmen, um existierende Standards nutzen zu können oder um gegebenenfalls anfallende Kosten für diese Schritte in den Antrag integrieren zu können.“ Online zugänglich unter: http://www.dfg.de/foerderung/antragstellung_begutachtung_entscheidung/antragstellende/antragstellung/index.html

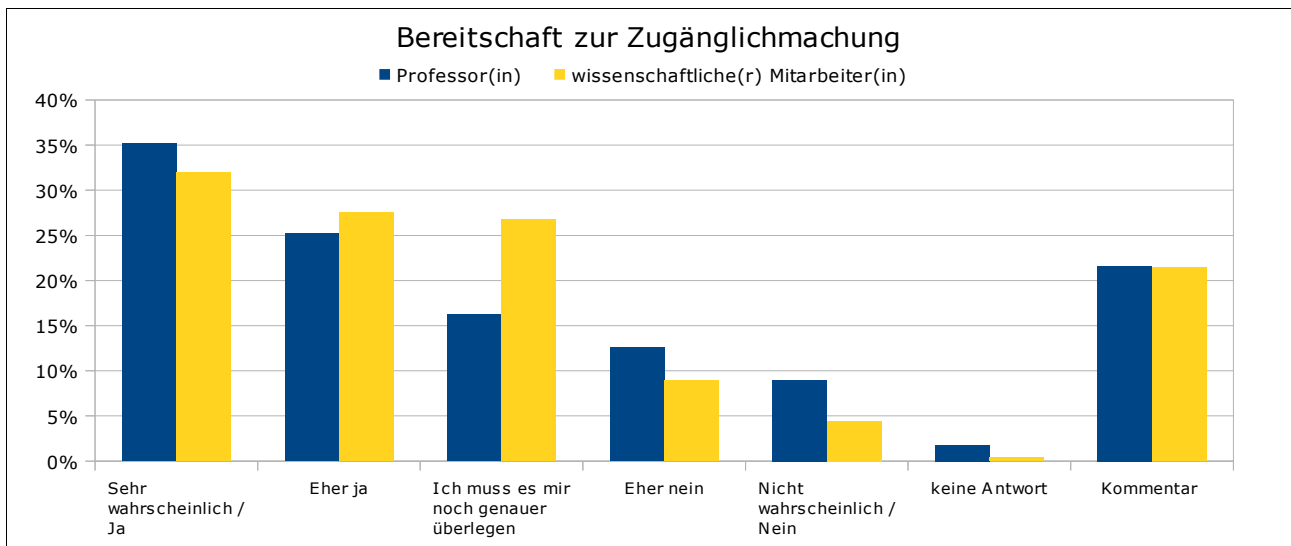


Abbildung 30: Bereitschaft zur Zugänglichmachung (Vergleich)

Aus den Ergebnissen geht hervor, dass wissenschaftliche Mitarbeiter der Zugänglichmachung von Forschungsdaten eher positiv gegenüber stehen, diese Möglichkeit aber oft genauer überlegen müssen. Das kann möglicherweise dadurch erklärt werden, dass sie die Verantwortung dafür nicht allein tragen bzw. dies zunächst mit leitenden Personen absprechen müssen (s. auch 3.4.1 Bereitschaft zur Zugänglichmachung).

Weiterhin wurde verglichen, welche Forschungsdaten die Teilnehmer in einem Datenarchiv ablegen würden. Die Ergebnisse werden in der Abbildung 31 dargestellt:

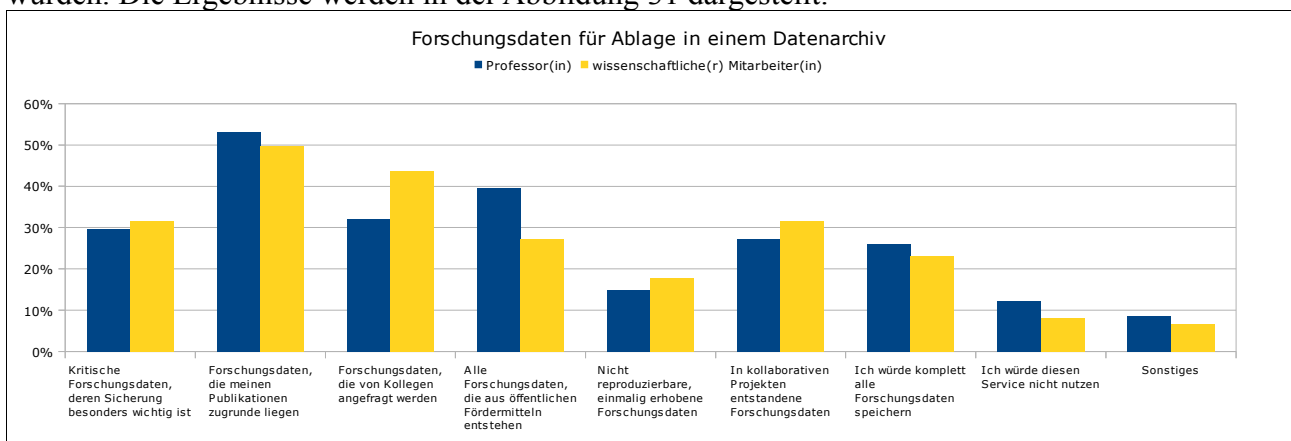


Abbildung 31: Forschungsdaten für Ablage in einem Datenarchiv (Vergleich)

Wenngleich bei dieser Frage keine großen Unterschiede festgestellt wurden, hat jeweils ein hoher Anteil von wissenschaftlichen Mitarbeitern „Forschungsdaten, die von Kollegen angefragt werden“ und ein hoher Anteil von Professoren „Alle Forschungsdaten, die aus öffentlichen Fördermitteln entstehen“ benannt. Dies kann mit einem unterschiedlich hohen Interesse am Austausch mit Kollegen bzw. einem höherem Bewusstsein bzgl. der Vorgaben der Drittmittelgeber in Zusammenhang stehen.

Daneben wurde verglichen, welches Datenarchiv die Teilnehmer für die Ablage ihrer Forschungsdaten bevorzugen würden. Die Ergebnisse werden in der Abbildung 32 dargestellt:

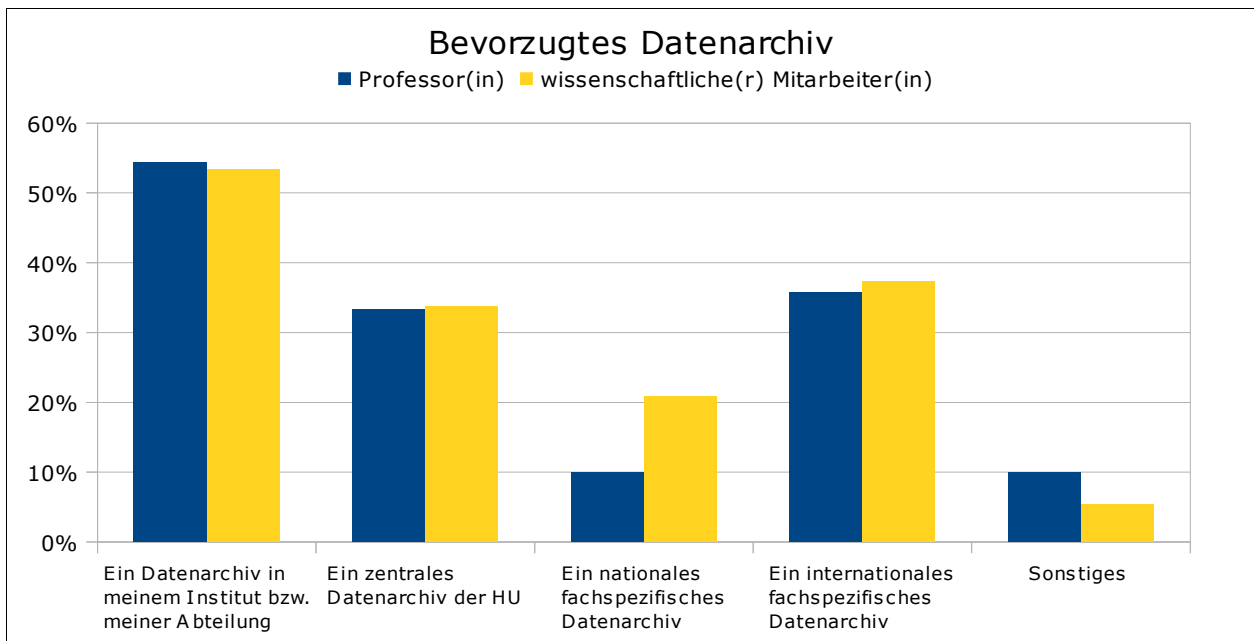


Abbildung 32: Bevorzugtes Datenarchiv (Vergleich)

Ähnlich wie bei der Frage zuvor zeigten sich keine wesentlichen Unterschiede zwischen den Teilnehmergruppen. Allein eine leichtere Präferenz für nationale fachspezifische Datenarchive war unter wissenschaftlichen Mitarbeitern zu sehen. Dies kann möglicherweise mit der persönlichen Karriereentwicklung verbunden sein.

Im Folgenden sollen hilfreiche Kriterien bei der Auswahl einer geeigneten Datenarchivs verglichen werden. Die Ergebnisse werden in den Abbildung 33a - e dargestellt:

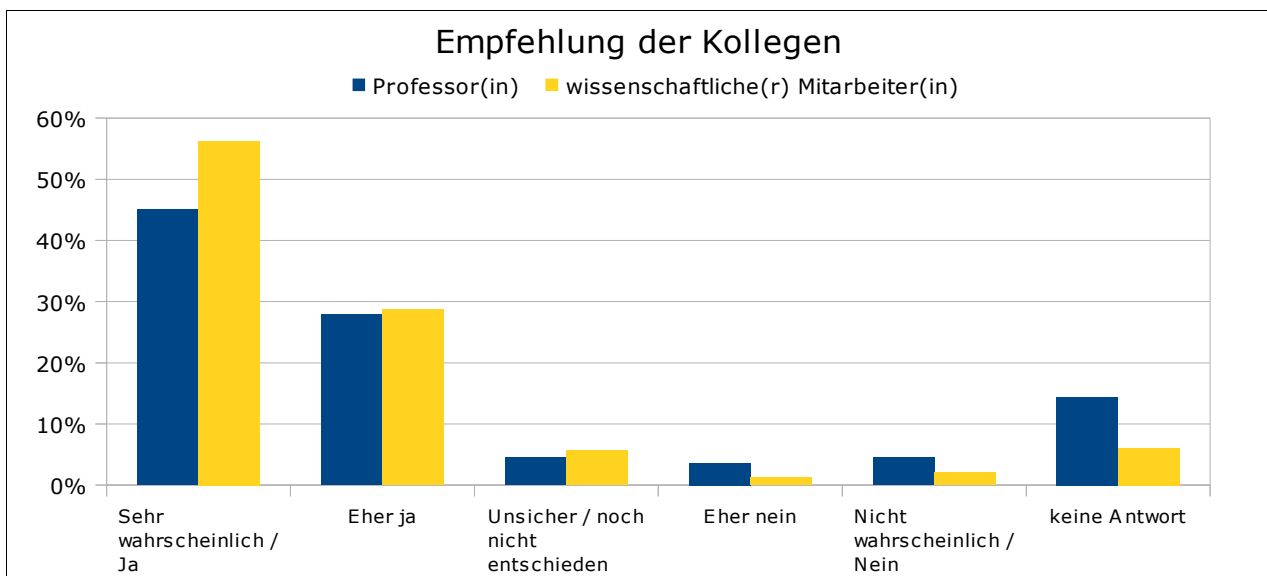


Abbildung 33a: Empfehlung der Kollegen (Vergleich)

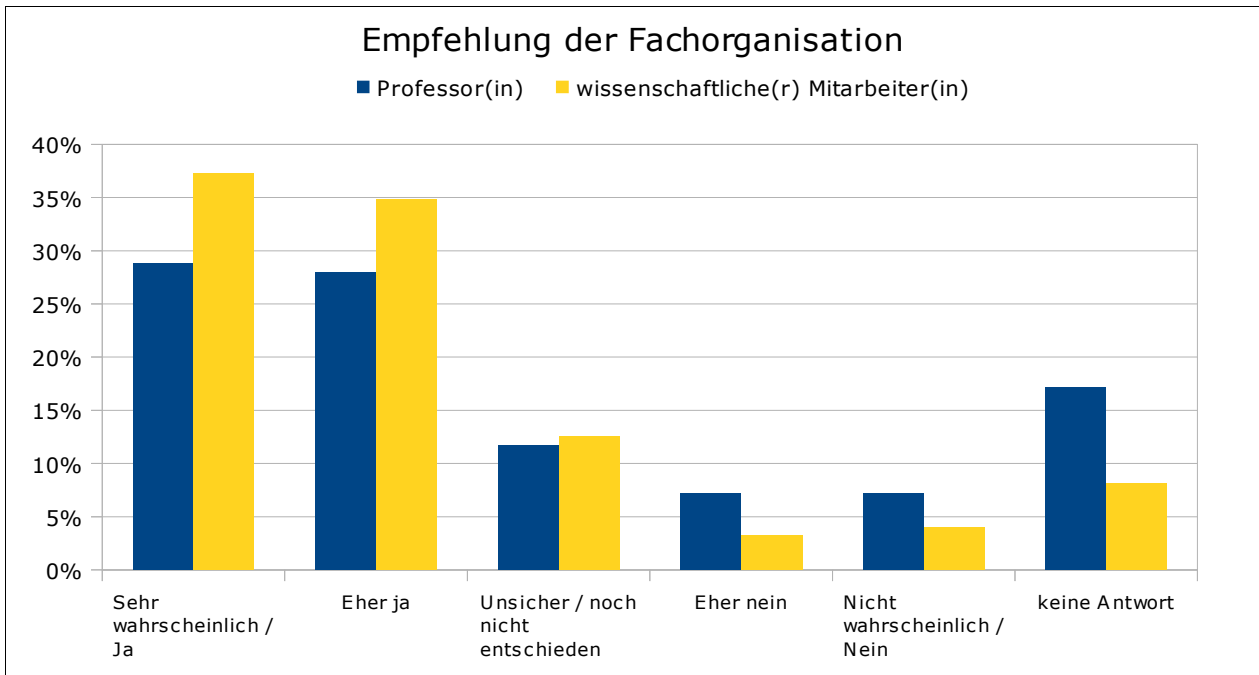


Abbildung 33b: Empfehlung der Fachorganisation (Vergleich)

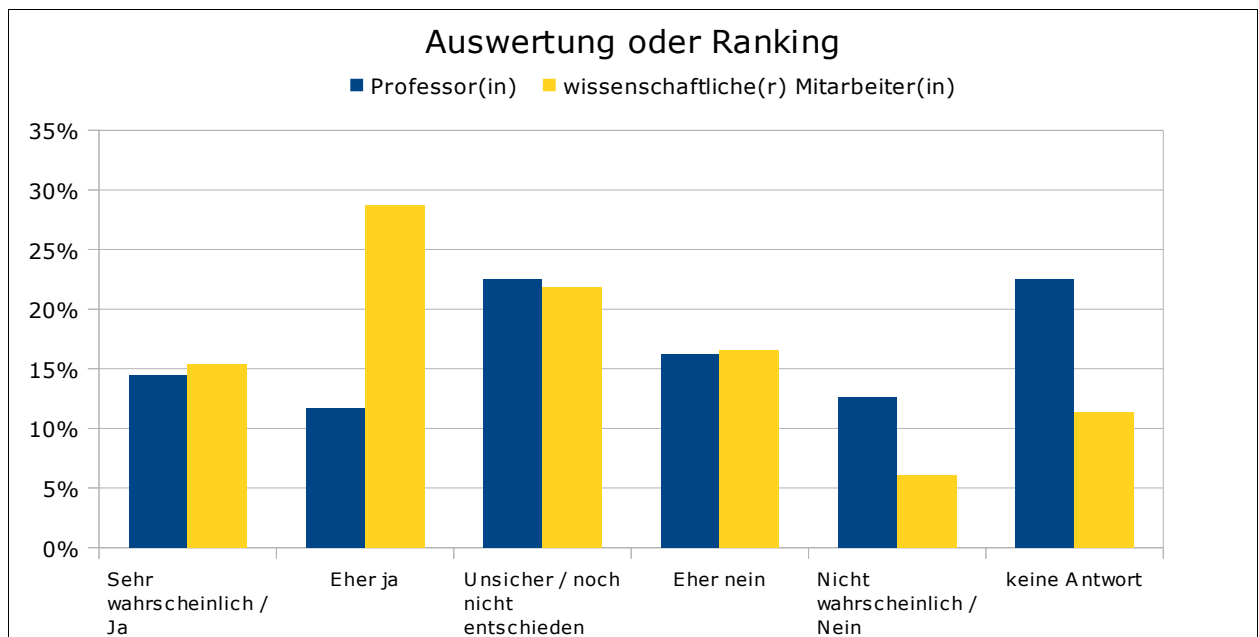


Abbildung 33c: Auswertung oder Ranking (Vergleich)

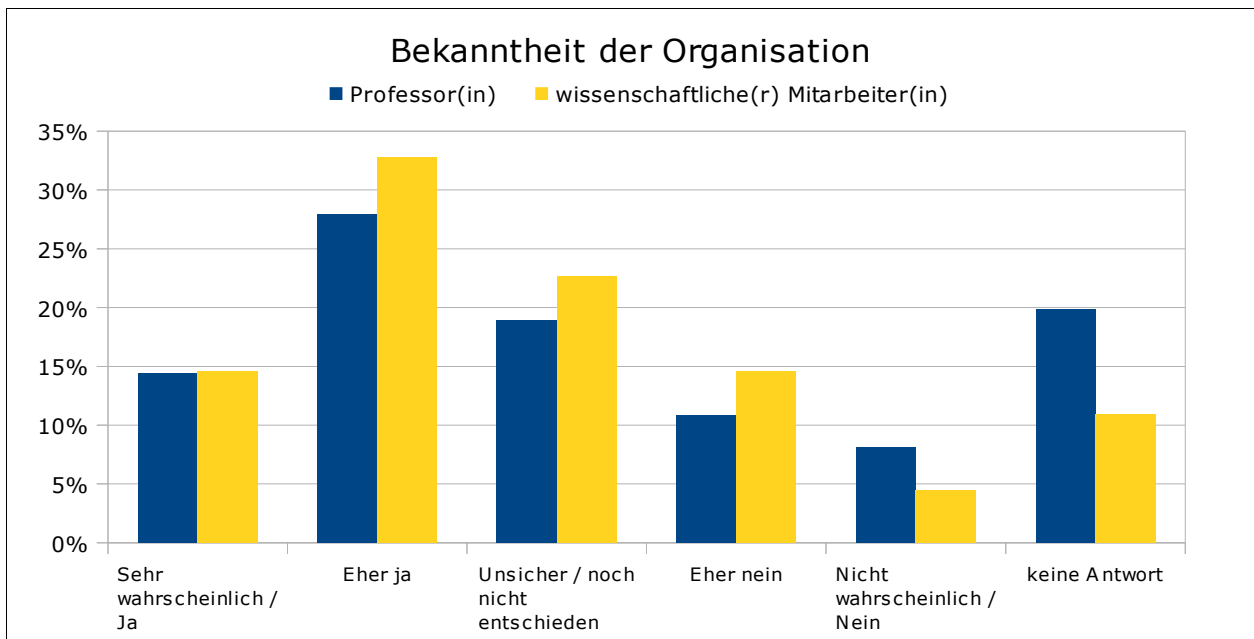


Abbildung 33d: Bekanntheit der Organisation (Vergleich)

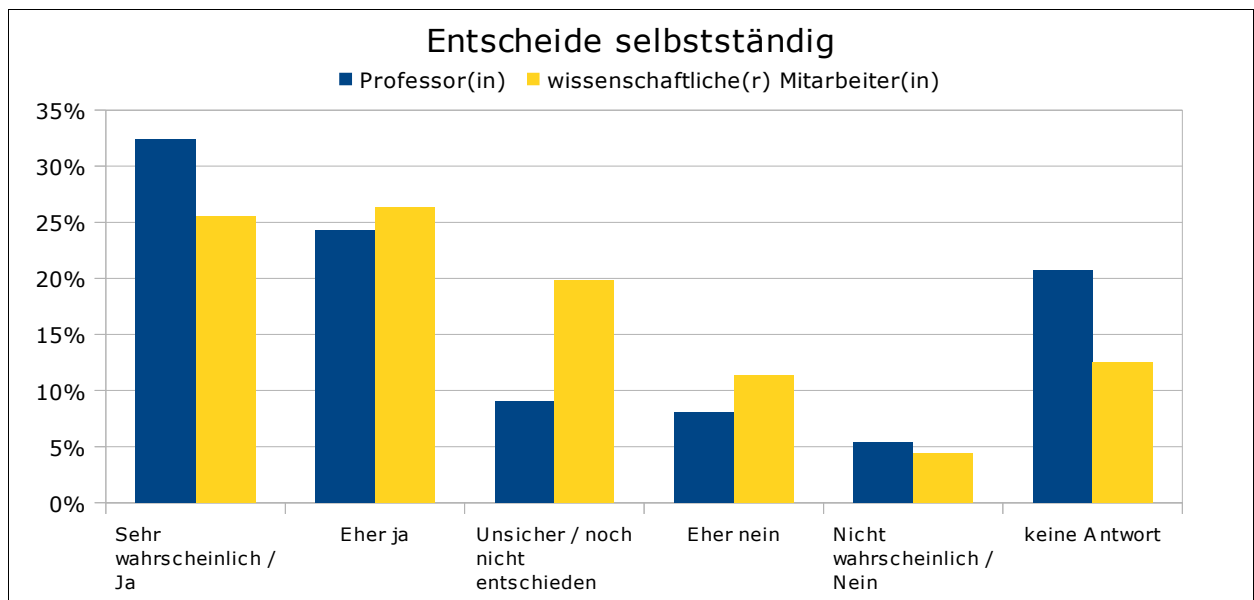


Abbildung 33e: Entscheide selbstständig (Vergleich)

Anhand der Ergebnissen lässt sich entnehmen, dass wissenschaftliche Mitarbeiter sich öfter auf die angebotenen Hilfskriterien verlassen würden bzw. seltener die Entscheidung über ein geeignetes Datenarchiv selbstständig treffen würden. Dies spricht möglicherweise dafür, dass die Professoren sich lieber auf ihre eigene Erfahrung verlassen, während wissenschaftliche Mitarbeiter offener für weitere Einflussfaktoren sind bzw. sie berücksichtigen.

Schließlich soll verglichen werden, ob sich die gewünschte Unterstützung zwischen den Teilnehmergruppen unterscheidet. Die Ergebnisse werden in der Abbildung 34 dargestellt:

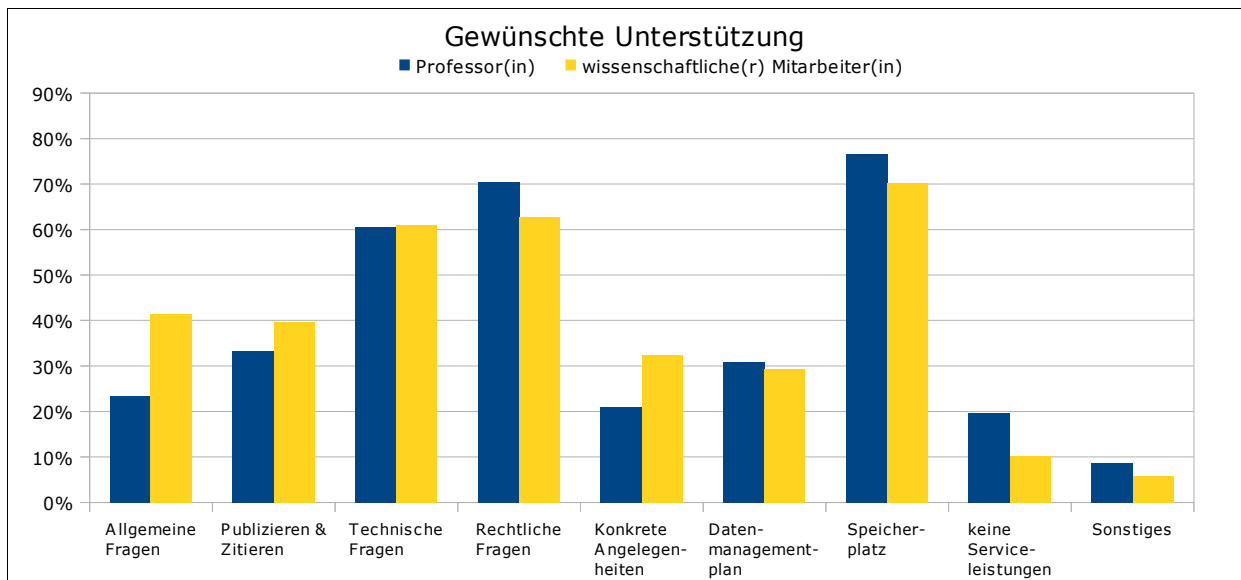


Abbildung 34: Gewünschte Unterstützung (Vergleich)

In der Abbildung ist eine fast gleichmäßige Verteilung der Antworten zu sehen. Bei genauerem Hinsehen zeigen sich jedoch einige Unterschiede. So wünschen sich wissenschaftliche Mitarbeiter eine Beratung und Unterstützung in allgemeinen Fragen zum Umgang mit Forschungsdaten und in konkreten Angelegenheiten vergleichsweise öfter als Professoren. Gleichzeitig wünschten sich die Professoren etwas häufiger Unterstützung hinsichtlich des Speicherplatzes und bei rechtlichen Fragen, waren aber auch häufiger unter denen, die keinen Bedarf an Serviceleistungen hatten. Dieses Ergebnis hängt möglicherweise mit einem größeren Informationsbedarf bei jüngeren Wissenschaftlern und der längeren Erfahrung bzw. der leitenden Funktion der Professoren zusammen.

4.4 Weiteres Interesse

Während das Interesse an einem Interview und an den Umfrageergebnissen bei Professoren relativ gleich war (22 % bzw. 24 %), war bei den wissenschaftlichen Mitarbeitern das Interesse an einem Interview (16 %) deutlich geringer als das Interesse an den Umfrageergebnissen (27 %). Ein solches eher „passives“ Interesse kann möglicherweise daraus folgen, dass die Entscheidungen bzw. die Verantwortung mit den leitenden Personen geteilt wird. Außerdem befanden sich einige Teilnehmer zum Zeitpunkt der Befragung erst am Anfang ihres Forschungsvorhabens und befürchteten demzufolge, in einem persönlichen Interview keine ausreichenden Erkenntnisse liefern zu können.

5 Zusammenfassung

Wissenschaftliche Einrichtungen wie Hochschulen und Forschungsinstitute sind zusammen mit den Förderorganisationen, Infrastrukturbetreibern und Fachgesellschaften die wichtigsten institutionellen Akteure im Betrieb der Forschung und der Verbreitung der wissenschaftlichen Erkenntnisse. Die Identifizierung des aktuellen Stands und der Anforderungen an künftige Serviceleistungen mithilfe einer Umfrage wird vom institutionellen Kontext der jeweiligen Einrichtung immer maßgeblich geprägt.

Was ist besonders an einer solchen Umfrage?

Die Gestaltung und Durchführung einer Umfrage zum Umgang mit Forschungsdaten an einer einzigen Hochschule mit zahlreichen Fachbereichen hat einen ganz anderen Charakter als eine Umfrage unter mehreren Einrichtungen europaweit oder Befragungen der Forscher aus einem konkreten Fachbereich. Dies ist vor allem mit den folgenden Aspekten verbunden:

- 1) **Multidisziplinäre Ausrichtung:** Im Vergleich zu fachspezifischen Einrichtungen sind die Hochschulen viel mehr multidisziplinär ausgerichtet. Das hat zur Folge, dass alle an einer Hochschule „beheimateten“ Fächer „gleichbehandelt“ werden müssen und keines davon zu sehr in den Mittelpunkt rücken darf. Bei der Gestaltung eines Fragebogens bedeutet dies, dass die Fragen und zugehörige Erläuterungen (z.B. Was sind Forschungsdaten?) mit einem gewissen Grad der Abstraktion formuliert werden müssen, um möglichst alle Fächer anzusprechen. Gleichzeitig entsteht dadurch das Risiko, dass die Aussagen sehr abstrakt oder undeutlich werden.
- 2) Berücksichtigung der **internen Organisationsstruktur:** Um eine hochschulweite Analyse initiieren zu können, wird die politische Unterstützung der Universitätsleitung (z.B. durch den Vizepräsidenten für Forschung oder Chief Information Officer) benötigt. Außerdem sollen relevante Dokumente (z.B. vom Akademischen Senat beschlossene Grundsätze zur Sicherung guter wissenschaftlicher Praxis) und bereits existierende Serviceangebote berücksichtigt bzw. als Antwortoptionen angeboten werden. Schließlich bedarf die Befragung der Mitarbeiter einer Hochschule der Abstimmung mit der zuständigen Stelle (z.B. dem/der Datenschutzbeauftragten und/oder dem Personalrat).
- 3) **Pragmatische Ausrichtung:** Die Durchführung einer solchen Umfrage an einer Hochschule ist unmittelbar mit der Frage verbunden, wie die Erkenntnisse daraus umgesetzt werden. Insbesondere können Erwartungen an den Ausbau der Unterstützungsangebote geweckt werden (z.B. Welche Serviceleistungen würden Sie sich wünschen?). Dies hat zur Folge, dass die Hochschulen den bestehenden Bedarf zur Kenntnis nehmen müssen und womöglich mit der „Lückenschließung“ beginnen sollten. Außerdem bietet sich die Möglichkeit, bereits existierende Serviceangebote und die jeweiligen Ansprechpartner unter den Forschern bekannter zu machen. Durch das Angebot eines persönlichen Interviews können zudem direkte Kontakte zu den Instituten und Arbeitsgruppen geknüpft werden und ein fundiertes Verständnis von deren typischen Arbeitsweisen und Anforderungen geschaffen werden.

Was sind die Erkenntnisse dieser Umfrage?⁴⁸

Durch eine vergleichsweise hohe Rücklaufquote und zahlreiche Kommentare hat sich gezeigt, dass die Fragen zum Umgang mit Forschungsdaten an der Humboldt-Universität zu Berlin aktuell und relevant sind. Insgesamt war der Rücklauf sowohl hoch, als auch breit über alle Fachbereiche

48 Vgl. Kindling et al. (2013a)

verteilt, was nicht zuletzt durch eine zusätzliche direkte Kontaktaufnahme mit weniger aktiven Instituten bzw. wissenschaftlichen Einrichtungen erreicht werden konnte. Wenngleich die Ergebnisse der Fachbereiche derzeit nicht miteinander verglichen wurden, stellen fachspezifische Unterschiede in der aktuellen Praxis eine weitere interessante Forschungsfrage dar. Im Folgenden wird auf die gewonnenen Erkenntnisse im Einzelnen eingegangen.

Ein wenig überraschend war zunächst die niedrige Teilnahmequote unter wissenschaftlichen Mitarbeitern (13 %) im Vergleich zu einer mehr als doppelt so hohen Teilnahmequote unter Professoren (29 %). Dies kann vermutlich daran liegen, dass der Anteil der Professoren im speziellen E-Mail-Verteiler für Projektleiter des Servicezentrums Forschung höher ist – über diesen Verteiler wurde die erste Einladung zur Teilnahme an der Umfrage verschickt. Zum anderen wäre auch vorstellbar, dass leitende Personen den Fragebogen für die gesamte Arbeitsgruppe ausgefüllt haben.

Die Ergebnisse zu den Quellen und Typen von Forschungsdaten bestätigten die Vermutung, dass die an der Humboldt-Universität vorhandenen Forschungsdaten sehr heterogen sind. Zudem waren unter meistgewählten Antwortoptionen die Textdokumente. Auch wenn es sich darüber streiten lässt, inwieweit diese als Forschungsdaten eingeordnet werden können, wurde eine fachübergreifende Verwendung in der Forschung deutlich. Weiterhin zeichneten sich unter fach- oder gerätespezifischen Daten statistische Auswertungen und Messreihen mit großem Abstand aus. Diese zwei Typen können daher ebenfalls als fachübergreifend gewertet werden.

Bei der Speicherung und Sicherung von Forschungsdaten wurden überwiegend lokale Optionen wie der dienstliche und private Rechner bzw. externe Festplatten und USB-Sticks bevorzugt. Allerdings wurden in Freitextkommentaren auch Cloud-Dienste genannt – mit über 50 Nennungen fiel insbesondere Dropbox auf. Dies signalisiert den Bedarf an flexiblen Filesystemen, mit deren Hilfe die Dateien mit Kooperationspartnern ortsunabhängig ausgetauscht werden können. Nicht zuletzt werden universitäre Angebote im Vergleich zu kommerziellen Anbietern durch ein höheres Vertrauen der Forschenden begünstigt.

Die vorherrschende eigene Verantwortung für Speicherung, Sicherung und Archivierung von Forschungsdaten zeigt zudem ein großes Potential für zentral organisierte Lösungen auf. Sollte zukünftig ein HU-weites Datenarchiv oder ein Forschungsdaten-Repository angeboten werden, so bieten die Ergebnisse zur Gesamtgröße der Forschungsdaten eine hilfreiche Orientierung zum benötigten Speicherplatz. Den Antworten zufolge schätzten die meisten Teilnehmer diesen auf 1-20 GB und 21-100 GB, so dass ein Wert von etwa 20-50 GB pro Nutzer bzw. ein Forschungsprojekt als notwendig angesehen werden könnte. Allerdings müsste zuvor geklärt werden, ob die Archivierung bzw. Veröffentlichung von Forschungsdaten nicht besser in fachspezifischen Repositorien erfolgen kann. Als Hilfestellung bietet sich das Verzeichnis für Forschungsdaten-Repositories [re3data.org](http://www.re3data.org) an.⁴⁹ Die Humboldt-Universität zu Berlin ist Projektpartner bei der Erstellung dieses Verzeichnisses.

Wenngleich die Ergebnisse zum Zugriff auf Forschungsdaten auf eine eher geschlossene Praxis hindeuten – die meisten Teilnehmer haben alleinig den Zugriff auf ihre Forschungsdaten oder machen diese nur der eigenen Arbeitsgruppe zugänglich – hatten ca. 20 % der Teilnehmer bereits von Erfahrungen in Veröffentlichung und Nachnutzung von Forschungsdaten berichtet (Kap. 3.3.2). Unter den konkreten Nennungen waren sowohl deutschsprachige als auch internationale Zeitschriftentitel und Datenarchive anzutreffen. Dies deutet auf die internationale Orientierung der Forschung hin und könnte u.a. für die Erstellung einer „Hochschulbibliographie“ oder einer anderen Form des Nachweises von den an der Humboldt-Universität produzierten Forschungsdaten hilfreiche Anhaltspunkte bieten.

Die an Kommentaren reichste Frage zur Sicherung guter wissenschaftlicher Praxis lieferte zugleich sehr wertvolle Einsichten zu problematischen Aspekten der Umsetzung dieser im Jahr 2002 an der

49 [re3data.org](http://www.re3data.org) – Registry of Research Data Repositories, <http://www.re3data.org/>

Humboldt-Universität beschlossenen Grundsätze. Schon die Verwendung verschiedener Fachbegriffe „Forschungsdaten“ (in der Umfrage) und „Primärdaten“ (in den Grundsätzen) sorgte für einige Verwirrung. Durch die dem Begriff „Primärdaten“ zugeschriebene empirische Arbeitsweise beurteilten manche Teilnehmer diese Grundsätze als nicht anwendbar für geisteswissenschaftliche Fachbereiche. Daneben wurden die Forschungsdaten oft aufbewahrt, ohne von den Grundsätzen zu wissen. Weiterhin wiesen die Teilnehmer darauf hin, dass die vorgegebene 10-Jahres-Frist für Sicherung und Aufbewahrung der Daten bei kurzen Arbeitsverträgen bzw. Verlassen der Forschungseinrichtung nur schwerlich eingehalten werden kann. Diese stichhaltigen Gründe machen es deutlich, dass zur Erfüllung der Regeln Guter wissenschaftlicher Praxis eine institutionelle Unterstützung notwendig ist.

Erstrebenswert ist ebenfalls, derartige Vorschriften auf den aktuellen Stand der informationswissenschaftlichen Diskussion zu bringen und zu harmonisieren. Damit z.B. auch geisteswissenschaftliche Wissenschaftsdisziplinen einbezogen werden, empfiehlt es sich, die Begrifflichkeiten umfassend zu definieren. Dies betrifft sowohl institutionelle Richtlinien als auch die von Drittmittelgebern. Insbesondere könnten an dieser Stelle die für viele wissenschaftliche Einrichtungen in Deutschland maßgeblichen „Vorschläge zur Sicherung guter wissenschaftlicher Praxis“⁵⁰ der DFG Vorschub leisten.⁵¹

Zudem hat sich erfreulicherweise gezeigt, dass die Mehrzahl der Teilnehmer auch auf freiwilliger Basis bzw. ohne institutionelle Verpflichtung zur Zugänglichmachung von Forschungsdaten bereit ist. Allerdings müssten dafür grundsätzliche Fragen (insb. zum Datenschutz und dem Urheberrecht) geklärt werden. Insbesondere der Bedarf nach einer rechtlichen Aufklärung wurde mehrmals thematisiert. Da derartige Fragen viele Einrichtungen vor große Herausforderungen stellen werden, scheint es angemessen, die Kräfte zusammenzuschließen und gemeinsam zu arbeiten (z.B. in Pilotprojekten oder einem Kompetenzzentrum am Beispiel des britischen DCC).

In Anlehnung an die Umfrageergebnisse kann die Schlussfolgerung gezogen werden, dass die Forschenden am ehesten bereit sind, sich für die Zugänglichmachung von Forschungsdaten zu entscheiden, wenn sie über drei Faktoren bestimmen können:

*{Welche Forschungsdaten?} * {Zu welchem Zeitpunkt?} * {Unter welchen Bedingungen?}*.

Angenommen, dass Forschungsdaten grundsätzlich öffentlich zugänglich gemacht werden sollen, können mithilfe dieser „Formel“ konkrete Kombinationen aus Typen von Forschungsdaten, dem richtigen Zeitpunkt und der geeigneten Publikationsweise identifiziert werden.⁵²

Schließlich wirkte sich im Vergleich der Antworten die längere Erfahrung der Professoren im Wissenschaftsbetrieb aus. Möglichkeiten und Regelungen zum Umgang mit Forschungsdaten waren ihnen öfter bekannt und sie machten von vorhandenen Unterstützungsangeboten mehr Gebrauch. Die Ergebnisse zeigen auch, dass die vielfältigen, damit verbundenen Kompetenzen effizienter an Nachwuchswissenschaftler weitergegeben werden müssen. In verschiedenen Studien zu dieser Thematik wird daher empfohlen, einen bewussten Umgang mit Forschungsdaten bereits in die Doktorandenausbildung zu integrieren.⁵³

Eine Diskussion der Ergebnisse und die Darstellung möglicher nächster Schritte erfolgt in Kindling et al. (2013a).

50 Darunter insb. Empfehlung 7 „Datenhaltung“, in: DFG, 1998

51 Siehe z.B. Research Councils UK Common Principles on Data Policy, die fachübergreifende Grundsätze zum Umgang mit Forschungsdaten im Vereinigten Königreich aufgestellt haben. Online zugänglich unter: <http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>

52 Das Zusammenspiel der hier genannten Faktoren bedarf einer genaueren Analyse, überschreitet jedoch den Rahmen des vorliegenden Umfrageberichts.

53 Vgl. Jones et al. (2013)

6 Literaturverzeichnis

Allianz der deutschen Wissenschaftsorganisationen (2010): *Grundsätze zum Umgang mit Forschungsdaten* Online zugänglich unter: <http://www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/grundsaeetze/> , zuletzt geprüft am 06.06.2013.

Burger, Marleen; Kindling, Maxi; Liebenau, Lisa; Lienhard, Claudia; Lilienthal, Svantje; Plewka, Paulina; Pohlkamp, Svenia; Reinhardt, Kristin; Rügenhagen, Melanie; Schulz, Kathleen; Simukovic, Elena; Sticht, Kendra; Walther, Marcel (2013): *Forschungsdatenmanagement an Hochschulen – Internationaler Überblick und Aspekte eines Konzepts für die Humboldt-Universität zu Berlin*. Version 1.1. edoc-Server der Humboldt-Universität zu Berlin. Online zugänglich unter: <urn:nbn:de:kobv:11-100210226>

Deutsche Forschungsgemeinschaft (1998): *Vorschläge zur Sicherung guter wissenschaftlicher Praxis. Empfehlungen der Kommission "Selbstkontrolle in der Wissenschaft"*. Denkschrift. Wiley-VCH Verlag, Weinheim. ISBN 3-527-27212-7. Online zugänglich unter: http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf , zuletzt geprüft am 06.06.2013.

Freiman, Lesley; Jones, Sarah; Molloy, Laura; Snow, Kellie; Ward, Catharine (2010): *Incremental Project Scoping study and implementation plan*. Online zugänglich unter: http://www.lib.cam.ac.uk/preservation/incremental/documents/Incremental_Scoping_Report_17091_0.pdf , zuletzt geprüft am 06.06.2013.

HATII at the University of Glasgow & Digital Curation Centre (2009): *Data Asset Framework Implementation Guide*. Online zugänglich unter: http://www.data-audit.eu/docs/DAF_Implementation_Guide.pdf , zuletzt geprüft am 06.06.2013.

Humboldt-Universität zu Berlin (2002): *Satzung über die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis und über den Umgang mit Vorwürfen wissenschaftlichen Fehlverhaltens*. Online zugänglich unter: <https://www.hu-berlin.de/forschung/services/0702gsgw.pdf> , zuletzt geprüft am 06.06.2013.

Jones, Sarah; Pryor, Graham; Whyte, Angus (2013): *How to Develop Research Data Management Services – a guide for HEIs*. DCC How-to Guides. Edinburgh: Digital Curation Centre. Online zugänglich unter: <http://www.dcc.ac.uk/resources/how-guides/how-develop-rdm-services> , zuletzt geprüft am 06.06.2013.

Kindling, Maxi (2013): *Doctoral theses' research data and metadata documentation*. Online zugänglich unter: <https://www.ibi.hu-berlin.de/forschung/publikationen/infomanagement/Kindling2013c> ; zuletzt geprüft am 15.10.2013.

Kindling, Maxi; Schirmbacher, Peter (2013): *Die digitale Forschungswelt als Gegenstand der Forschung*. In: Information - Wissenschaft & Praxis Bd. 64, H. 2-3, S. 137-148. Online zugänglich unter: <10.1515/iwp-2013-0017>

Kindling, Maxi; Schirmbacher, Peter; Simukovic, Elena (2013a): *Forschungsdatenmanagement an Hochschulen: das Beispiel der Humboldt-Universität zu Berlin*. In: LIBREAS. Library Ideas, 23, S.

43-63. Online zugänglich unter: [urn:nbn:de:kobv:11-100212700](https://nbn-resolving.org/urn:nbn:de:kobv:11-100212700)

Kindling, Maxi; Schirnbacher, Peter; Simukovic, Elena (2013b): „*It's all about data*“. *Digitale Forschungsdaten an der Humboldt-Universität zu Berlin*. Vortrag beim Berliner Bibliothekswissenschaftlichen Kolloquium am 21.05.2013. Online zugänglich unter: https://www.cms.hu-berlin.de/dataman/bbk-pdf/at_download/file , zuletzt geprüft am 06.06.2013.

Kuipers, Tom; van der Hoeven, Jeffrey (2009): *PARSE.Insight: Insight into digital preservation of research output in Europe*. Survey report. Deliverable D3.4. Online zugänglich unter: http://www.parse-insight.eu/downloads/PARSE-Insight_D3-4_SurveyReport_final_hq.pdf , zuletzt geprüft am 06.06.2013.

Porsche, Jana (2012): *Actual state of research data @ ISTAustria*. Project Report. Version 1.0. IST Austria. IST Austria Repository. Online zugänglich unter: <https://repository.ist.ac.at/103/> , zuletzt geprüft am 06.06.2013.

Porsche, Jana (2013): *Institutional Repository für Forschungsdaten - Unterstützung für Wissenschaft?!*. Vortrag beim 5. Kongress Bibliothek & Information Deutschland, 11.-14. März 2013 in Leipzig. OPUS-Publikationsserver des Berufsverbands Information Bibliothek e.V. Online zugänglich unter: [urn:nbn:de:0290-opus-13978](https://nbn-resolving.org/urn:nbn:de:0290-opus-13978)

Quandt, Markus (2012): *Ergebnisse der Befragung der Institute der Leibniz-Gemeinschaft. Leibniz-Forschungsdatenumfrage 2010/2011*. Vortrag beim Workshop des AK Forschungsdaten der Leibniz-Gemeinschaft am 10.05.2012 in Berlin. Online zugänglich unter: <http://www.tib-hannover.de/fileadmin/aktuell/ak-forschungsdaten/Leibniz-Forschungsdatenumfrage-Markus-Quandt-GESIS.pdf> , zuletzt geprüft am 25.06.2013.

Quirk, Rachel; Olver, Martin; Hammond, Max; Davies, Claire (2008): *The Guide to Researching Audiences*. Version 1.1. The Strategic Content Alliance. Online zugänglich unter: <http://www.jisc.ac.uk/media/documents/publications/general/2009/scaaudiencesfullguide.pdf> , zuletzt geprüft am 06.06.2013.

Simukovic, Elena (2012): *Das Erdöl von morgen*. In: HUMBOLDT - Die Zeitung der Alma Mater Berolinensis, Ausgabe 2 vom 15. November 2012, S. 7. Online zugänglich unter: https://www.hu-berlin.de/pr/medien/publikationen/humboldt/2013/201211/humboldt_201211.pdf , zuletzt geprüft am 06.06.2013.

Töwe, Matthias; Scheid, Susanne (2011): *User Expectations in Archived Research Data*. Digital Preservation Summit 2011, Hamburg. Online zugänglich unter: <http://www.digitalpreservationsummit.de/presentations/toewe.pdf> , zuletzt geprüft am 30.05.2013.

Vlaeminck, Sven (2008): *AP 2 - Erfassen und Kategorisieren relevanter Datenbestände*. KoLaWiss Report R2 - AP2 „Kategorisierung“. Version 1.2. Online zugänglich unter: http://kolawiss.uni-goettingen.de/projektergebnisse/AP2_Report.pdf , zuletzt geprüft am 06.06.2013.

Zugrundeliegende Forschungsdaten:

Simukovic, Elena; Kindling, Maxi; Schirnbacher, Peter (2013): *Ergebnisse der Umfrage zum Umgang mit digitalen Forschungsdaten an der Humboldt-Universität zu Berlin*. ZENODO-Repository. Online zugänglich unter DOI: [10.5281/zenodo.7446](https://doi.org/10.5281/zenodo.7446)
Dateiname: Umfrage_Forschungsdaten_2013_HU-Berlin.ods

Simukovic, Elena; Kindling, Maxi; Schirnbacher, Peter (2013): *Ergebnisse der Umfrage zum Umgang mit digitalen Forschungsdaten an der Humboldt-Universität zu Berlin, Vergleich der Antworten zwischen Teilnehmergruppen "Professor(in)" und "wissenschaftliche(r) Mitarbeiter(in)"*. ZENODO-Repository. Online zugänglich unter DOI: [10.5281/zenodo.7447](https://doi.org/10.5281/zenodo.7447)
Dateiname: Umfrage_Forschungsdaten_2013-Vergleich_nach_Groupen.ods

7 Anhänge

7.1 Einladungstext (Deutsche Version)

Einladung zur Mitgestaltung des Forschungsdatenmanagements an der HU

Liebe Forscherinnen und Forscher,
liebe Kolleginnen und Kollegen,

die Forschungsabteilung, die Universitätsbibliothek und der Computer- und Medienservice arbeiten seit einiger Zeit zusammen an der Entwicklung eines Konzeptes für das Forschungsdatenmanagement an der HU.

Mit dieser Initiative möchten wir Sie bei der Umsetzung der Regeln Guter wissenschaftlicher Praxis besser unterstützen und auf die aussichtsreichen neuen Möglichkeiten aufmerksam machen.

Damit wir den vielfältigen Anforderungen verschiedener Forschergruppen gerecht werden können, laden wir Sie herzlich ein, an einer universitätsweiten Umfrage teilzunehmen und von den gängigen Praktiken in Ihrer Arbeitsgruppe oder Ihrem Fachbereich zu berichten.

Die Teilnahme ist freiwillig und anonym. Ihre Antworten werden der Gestaltung zentraler Serviceleistungen dienen.

Die Umfrage können Sie bis einschließlich 1. März 2013⁵⁴ unter <https://umfrage.hu-berlin.de/index.php?sid=68656&lang=de> erreichen.

Bitte teilen Sie den Link auch Ihren Kolleginnen und Kollegen mit, damit der jeweilige Fachbereich zureichend repräsentiert wird.

Wir freuen uns auf Ihre Teilnahme und stehen für Rückmeldungen jederzeit zur Verfügung.

Mit den besten Grüßen

Im Namen der Initiative
i.A. Elena Simukovic

Nähere Informationen finden Sie unter <http://www.cms.hu-berlin.de/dataman>

⁵⁴ Die Laufzeit der Umfrage wurde später bis 8. März 2013 verlängert.

7.2 Fragebogen (Deutsche Version)

Fachbereich und Position

Angaben zum Institut bzw. Forschungszentrum und der Position

1 Bitte wählen Sie das Institut bzw. das Forschungszentrum aus, bei dem Sie hauptsächlich beschäftigt sind. *

Bitte wählen Sie nur eine der folgenden Antworten aus:

[Liste der Einrichtungen zur Auswahl]

2 Ihre Position an dem oben ausgewählten Institut bzw. Forschungszentrum:

Bitte wählen Sie alle zutreffenden Antworten aus:

- Professor(in)
- Leiter(in) der Arbeitsgruppe
- Wissenschaftliche(r) Mitarbeiter(in)
- Doktorand(in)
- IT-Mitarbeiter(in)
- Studentische(r) Mitarbeiter(in)
- Sonstiges: _____

Falls nicht zutreffend oder zum Ergänzen nutzen Sie bitte das Freitext-Feld.

Datentypen

Charakteristik, Größe, Formate

3 Aus welchen Quellen schöpfen Sie Ihre Forschungsdaten hauptsächlich?

Bitte wählen Sie alle zutreffenden Antworten aus:

- Beobachtungen
- Experimente
- Simulationen
- Abbildungen von Objekten
- Umfragen und Interviews
- Statistik und Referenzdaten
- Logfiles und Nutzungsdaten
- Textdokumente
- Sonstiges: _____

Falls nicht zutreffend oder zum Ergänzen nutzen Sie bitte das Freitext-Feld.

4 Um welche Datentypen handelt es sich? *

Bitte wählen Sie alle zutreffenden Antworten aus:

- Bilder
- mehrdimensionale Visualisierungen oder Modelle
- Audio-Aufzeichnungen
- Video-Aufzeichnungen
- Texte
- Tabellen
- Datenbanken
- Programme und Anwendungen
- Fach- oder gerätespezifische Daten

- Sonstiges: _____

Für detailliertere Datentypen wählen Sie bitte "Fach- oder gerätespezifische Daten".

5 Um welche Art fach- oder gerätespezifische Daten handelt es sich? *

Bitte wählen Sie alle zutreffenden Antworten aus:

- Statistische Auswertung
- Text-Korpora / Annotationen
- GIS-Daten
- Topographische Daten
- Satellitenbilder
- Fernerkundungsdaten
- Spektre
- Messreihen
- Klimamodellierungen
- Patienten-Daten
- Fragebögen
- Sonstiges: _____

6 Wo speichern Sie Ihre Forschungsdaten? *

Bitte wählen Sie alle zutreffenden Antworten aus:

- lokal auf meinem privaten Rechner
- lokal auf meinem dienstlichen Rechner
- zentral auf dem Server im Institut
- zentral auf dem Server der HU
- bei einem externen Datenarchiv einer wissenschaftlichen Einrichtung
- bei einem externen Datenarchiv eines kommerziellen Anbieters
- Sonstiges: _____

Als "Datenarchive" werden Online-Dienste betrachtet, wo Sie Ihre Forschungsdaten speichern und für andere zugänglich machen können.

7 Auf welchen Datenträgern speichern Sie Ihre Forschungsdaten zusätzlich?

Bitte wählen Sie alle zutreffenden Antworten aus:

- externe Festplatte
- USB-Stick
- CD, DVD
- Server der Einrichtung
- auf keinem Datenträger
- Sonstiges: _____

8 Bitte schätzen Sie die Gesamtgröße Ihrer Forschungsdaten bezogen auf den benötigten Speicherplatz.

Bitte wählen Sie nur eine der folgenden Antworten aus:

- < 1 GB
- 1 - 20 GB
- 21 - 100 GB
- 101 GB - 1 TB
- > 1 TB
- Ich kann es nicht einschätzen.

9 Wie oft erstellen Sie bzw. werden die Sicherungskopien erstellt?

Bitte wählen Sie nur eine der folgenden Antworten aus:

- täglich

- wöchentlich
- monatlich
- vierteljährlich
- jährlich
- unregelmäßig / ad hoc
- abhängig vom Dienstleister bzw. Server-Betreiber
- nie
- Sonstiges: _____

10 Wer ist verantwortlich für die Speicherung, Sicherung oder Archivierung Ihrer Forschungsdaten?

Bitte wählen Sie alle zutreffenden Antworten aus:

- Ich selbst
- Mein(e) Assistent(in)
- Doktorand(in) am Lehrstuhl oder Institut
- Projekt- / Gruppenleiter(in)
- Spezielles Personal innerhalb der Arbeitsgruppe
- Personal am CMS
- Personal an der (Zweig-)Bibliothek
- Externer Dienstleister
- Sonstiges: _____

Zugang

Rechte, Zugänglichkeit, Einschränkungen

11 Wer außer Ihnen darf bzw. kann auf Ihre Forschungsdaten zugreifen? *

Bitte wählen Sie alle zutreffenden Antworten aus:

- Meine Arbeitsgruppe
- Mein Institut
- HU insgesamt
- (Fach)öffentlichkeit
- Forschungsdaten werden den Interessenten per Anfrage bereitgestellt
- Nur ich selbst
- Sonstiges: _____

12 Haben Sie schon einmal...

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	Ja.	Nein, aber ich habe es vor.	Nein, diese Möglichkeit war mir nicht bekannt.	Nein, in der nächsten Zeit habe ich es nicht vor.
Ihre Forschungsdaten in einem Datenarchiv abgelegt?				
Forschungsdaten anderer Forscher von einem Datenarchiv heruntergeladen bzw. zitiert?				
Ihre Forschungsdaten mit dem Manuskript bei einer Zeitschrift eingereicht bzw. nachgewiesen?				

Als "Datenarchive" werden Online-Dienste betrachtet, wo Sie Ihre Forschungsdaten speichern und für andere zugänglich machen können.

13 Bitte geben Sie den Namen oder die URL des obengenannten Datenarchivs bzw. der Zeitschrift an.

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

----- Scenario -----

Antwort war 'Nein, aber ich habe es vor.' oder 'Ja.' bei Frage '12' (Haben Sie schon einmal...)

Bitte geben Sie Ihre Antwort hier ein: _____

14 In den Grundsätzen zur Sicherung guter wissenschaftlicher Praxis der HU (2002) heißt es: "Primärdaten als Grundlagen für Veröffentlichungen sollen auf haltbaren und gesicherten Trägern in der Arbeitsgruppe/ Einrichtung, wo sie entstanden sind, für zehn Jahre zugänglich bleiben." Berücksichtigen Sie diese Grundsätze?

Bitte wählen Sie nur eine der folgenden Antworten aus:

- Ja, die Daten werden für zehn Jahre oder mehr aufbewahrt.
- Die Daten werden nur während oder kurz nach dem Abschluss des Projektes aufbewahrt.
- Nein, diese Grundsätze kenne ich nicht.
- Mir ist die genaue Situation nicht bekannt.

Bitte schreiben Sie einen Kommentar zu Ihrer Auswahl: _____

Die Begriffe "Primärdaten" und "Forschungsdaten" werden in diesem Zusammenhang synonym verwendet. Wenn möglich, beschreiben Sie bitte kurz Ihre Erfahrung im Kommentar.

15 Gibt es weitere formale Richtlinien oder Prozeduren, die den Umgang mit Forschungsdaten genauer bestimmen?

Bitte wählen Sie alle zutreffenden Antworten aus:

- Richtlinien meiner Arbeitsgruppe oder Abteilung.
- Richtlinien meiner Fachorganisation.
- Richtlinien des Drittmittelgebers (z.B. DFG, BMBF, EU oder einer Stiftung).
- Mir ist die genaue Situation nicht bekannt.
- Sonstiges: _____

Wenn möglich, machen Sie bitte auch nähere Angaben.

Infrastruktur und Service

Aktuelle und gewünschte Infrastruktur

16 Wären Sie grundsätzlich bereit, bestimmte Forschungsdaten in einem Datenarchiv abzulegen bzw. zugänglich zu machen? *

Bitte wählen Sie nur eine der folgenden Antworten aus:

- Sehr wahrscheinlich / Ja.
- Eher ja.
- Ich muss es mir noch genauer überlegen.
- Eher nein.
- Nicht wahrscheinlich / Nein.

Bitte schreiben Sie einen Kommentar zu Ihrer Auswahl: _____

Wenn möglich, erläutern Sie bitte Ihre Meinung näher im Kommentar-Feld.

17 Welche Forschungsdaten würden Sie dort ablegen?

Bitte wählen Sie alle zutreffenden Antworten aus:

- Kritische Forschungsdaten, deren Sicherung besonders wichtig ist.
- Forschungsdaten, die meinen Publikationen zugrunde liegen.

- Forschungsdaten, die von Kollegen angefragt werden.
- Alle Forschungsdaten, die aus öffentlichen Fördermitteln entstehen.
- Nicht reproduzierbare, einmalig erhobene Forschungsdaten.
- In kollaborativen Projekten entstandene Forschungsdaten.
- Ich würde komplett alle Forschungsdaten speichern.
- Ich würde diesen Service nicht nutzen.
- Sonstiges: _____

18 Welches Datenarchiv würden Sie dafür am ehesten nutzen?

Bitte wählen Sie alle zutreffenden Antworten aus:

- Ein Datenarchiv in meinem Institut bzw. meiner Abteilung.
- Ein internationales fachspezifisches Datenarchiv.
- Ein nationales fachspezifisches Datenarchiv.
- Ein zentrales Datenarchiv der HU.
- Sonstiges: _____

19 Welche Kriterien würden Ihnen helfen, ein geeignetes Datenarchiv auszuwählen?

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	Sehr wahrscheinlich / Ja.	Eher ja.	Unsicher / noch nicht entschieden.	Eher nein.	Nicht wahrscheinlich / Nein.
Empfehlung meiner Kollegen					
Empfehlung meiner Fachorganisation					
Auswertung oder Ranking in einem unabhängigen Verzeichnis					
Bekanntheit der Organisation, die das Datenarchiv betreibt					
Am besten entscheide ich es selbstständig					
Sonstiges (bitte genauer angeben)					

20 Bitte beschreiben Sie sonstige Kriterien im Freitext-Feld.

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

----- Scenario -----

Antwort war [alle Antwortoptionen] bei Frage '19' (Welche Kriterien würden Ihnen helfen, ein geeignetes Datenarchiv auszuwählen?) und Antwort war 'Sonstiges (bitte genauer angeben)'.

Bitte geben Sie Ihre Antwort hier ein: _____

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

Bitte wählen Sie nur eine der folgenden Antworten aus:

- Ja, regelmäßig.
- Ja, solche Kooperationen entstehen ab und zu.
- Ja, aber eher in Ausnahmefällen.
- Nein, solche Kooperationen gibt es in meinem Arbeitsbereich nicht.

Bitte schreiben Sie einen Kommentar zu Ihrer Auswahl: _____

Wenn möglich, beschreiben Sie bitte kurz die gängige Praxis Ihrem Fachbereich.

22 Welche Serviceleistungen würden Sie sich von der HU wünschen?

Bitte wählen Sie alle zutreffenden Antworten aus:

- Beratung zum Publizieren und Zitieren von Forschungsdaten.
- Beratung in allgemeinen Fragen zum Umgang mit Forschungsdaten.
- Beratung bei technischen Fragen (z.B. Metadaten, Standards, langfristige Archivierung).
- Beratung bei rechtlichen Fragen (z.B. Zugangseinschränkungen, Umgang mit sensiblen Daten, Nutzung von Lizenzen).
- Unterstützung bei konkreten Angelegenheiten (z.B. beim Einreichen eines Manuskripts in einer Zeitschrift, die Veröffentlichung von Forschungsdaten vorgibt).
- Unterstützung beim Erstellen eines Datenmanagementplans, wenn ein solcher von der Förderorganisation angefordert wird.
- Einen gesicherten und verlässlichen Speicherplatz für meine Forschungsdaten.
- Ich brauche keine Serviceleistungen.
- Sonstiges: _____

Spezielle Wünsche oder Anregungen können gerne auch im persönlichen Gespräch mitgeteilt werden.

Interesse

Interesse am Interview oder Umfrageergebnissen

23 Wären Sie bereit, von Ihren Erfahrungen zum Umgang mit Forschungsdaten und Erwartungen an zentrale Serviceleistungen in einem persönlichen Gespräch zu berichten? *

Bitte wählen Sie nur eine der folgenden Antworten aus:

- Nein.
- Ja: _____

Wenn ja, bitte geben Sie Ihre Kontaktdaten an (bevorzugt E-Mail-Adresse).

24 Wenn Sie Interesse an Ergebnissen der Umfrage haben, können Sie hier Ihre E-Mail-Adresse angeben.

Bitte geben Sie Ihre Antwort hier ein: _____

Angabe ist freiwillig.