

## 1 Zusammenfassung

## 2 Einführung

## 3 Methodenbericht

Der Fragebogen wurde aus der Expertise der Mitglieder des AK Forschungsdaten heraus entwickelt, um die Situation in den Instituten der Leibniz-Gemeinschaft hinsichtlich Forschungsdatenmanagement und ‚data sharing‘ im Überblick beurteilen zu können.

Die Umfrage wurde schriftlich als Online-Erhebung durchgeführt, d.h., die Befragten beantworteten nach einem passwortgeschützten Login die Fragen mithilfe ihres Web-Browsers am eigenen Arbeitsplatz. Die organisatorische Durchführung erfolgte durch das GESIS Leibniz-Institut für Sozialwissenschaften als Mitglied des AK Forschungsdaten. Mit Unterstützung des Leibniz-Präsidiums wurden Anschreiben per Post und E-Mail zunächst an die Institutsleiter versandt. Die Rekrutierung für die Teilnahme sah auf einer nächsten Stufe die Selbst-Registrierung von einer oder mehreren Personen im Online-Umfragesystem vor, die für ihr Institut bzw. einzelne Abteilungen davon die Fragen beantworten sollten. An diese Personen wurden gegebenenfalls nach Beginn der Erhebungsphase mehrfach per E-Mail Erinnerungsnachrichten versandt.

Nachdem in einer ersten Erhebungsphase vom 26.5.-7.7.2010 die Teilnahme von 52 Instituten erreicht wurde, die Ausschöpfung aber nicht in allen Sektionen befriedigend war, wurde auf Beschluss des Leibniz-Präsidiums eine Nacherhebung durch den AK Forschungsdaten durchgeführt. Diese zweite Welle war vom 15.3.-21.4.2011 im Feld und hat mit identischen Methoden, jedoch einem veränderten Anschreiben an die Institute (es konnte nun ausdrücklich auf den Beschluss des Leibniz-Präsidiums verwiesen werden) eine insgesamt sehr gute Teilnahmequote der Institute ergeben.

Nachdem in der zweite Erhebungswelle 25 weitere Institute teilgenommen haben, sind die Sektionen der Leibniz-Gemeinschaft alle mit mindestens 75 % der Institute repräsentiert, in Sektion E haben mit sogar alle zehn Institute der Sektion die Umfrage beantwortet. Damit haben 77 von insgesamt 89 Leibniz-Instituten teilgenommen (inkl. assoz. Institute). Weil in 22 Instituten die Option wahrgenommen wurde, für unterschiedliche Abteilungen des Institutes jeweils einen eigenen Fragebogen auszufüllen, liegen insgesamt 136 auswertbare Fragebögen vor. Die Beantwortung durch mehrere Abteilungen je Institut kam in den Sektionen B und C häufig vor, in den Abteilungen D und E hingegen nur jeweils einmal.<sup>1</sup>

---

<sup>1</sup> Abteilungsamen wurden leider nicht immer angegeben, jedoch wurde in allen Fällen mit fehlenden Abteilungsangaben eine Plausibilitätsprüfung über die inhaltlichen Antwortmuster gemacht. Danach handelt es sich bei diesen Fällen nicht um Dubletten.

Sektion	Teilnehmende Institute	Gesamtzahl Antworten	Gesamtzahl Institute (mit assoz.) in der Sektion	Anteil erfasster Institute an allen Instituten
A - Geisteswissenschaften und Bildungsforschung	14	19	15	93,3
B - Wirtschafts- und Sozialwissenschaften, Raumwissenschaften	17	31	18	94,4
C - Lebenswissenschaften	18	54	24	75,0
D - Mathematik, Natur- und Ingenieurwissenschaften	18	20	22	81,8
E - Umweltwissenschaften	10	12	10	100,0
Gesamt	77	136	89	86,5

## 4 Auswertungen

Die Tatsache, dass viele Institute die Option gewählt haben, nur einen Fragebogen für das gesamte Institut auszufüllen, während andere Institute für verschiedene Abteilungen jeweils einen eigenen Fragebogen ausgefüllt haben, muss bei allen Auswertungen berücksichtigt werden. Ein großer Teil der unten berichteten Auswertungen bezieht sich ausschließlich auf die Institutebene (mit 77 Fällen). Das bedeutet, dass Antworten von mehreren Abteilungen jeweils desselben Institutes soweit möglich auf Institutebene aggregiert werden mussten.<sup>2</sup> Wenn eine Aggregation der Auswertungen auf Institutebene nicht notwendig oder nicht möglich ist, basieren die Angaben auf allen 136 Fällen. Dies kann dann sinnvoll sein, wenn das Hauptinteresse nicht darauf liegt, die Institute der Leibniz-Gemeinschaft in ihrer Gesamtheit abzubilden, sondern wenn es darum geht, die Urteile möglichst vieler Befragter als ‚Expertinnen‘ und ‚Experten‘ für ihr jeweiliges Feld direkt zu betrachten. Solche Auswertungen auf Abteilungs- bzw. Befragtenebene werden separat gekennzeichnet.

Die Perspektive dieses Berichts ist zunächst, die Situation in der Leibniz-Gemeinschaft zu beschreiben. Daher werden überwiegend einfache Anteilswerte, absolute Zahlen oder Durchschnitte berichtet. Sofern die abgefragten Sachverhalte auf hinreichend viele Institute zutrafen, wird ferner als allererster Zugriff auf eine disziplin-spezifische Beschreibung die Sichtweise über die unterschiedlichen Sektionen der Leibniz-Gemeinschaft gewählt.

Zum vollständigen Verständnis der berichteten Tabellen empfiehlt es sich, den Fragebogen mit den vollständigen Frageformulierungen zur Hand zu haben. Diese werden im vorliegenden internen Bericht in der Regel nicht wiedergegeben.

### 4.1 Forschungsdaten in der Disziplin

#### F.1, Datentypen

Einleitend wurde nach der Art der Forschungsdaten gefragt, die in der Disziplin typischerweise verwendet werden. Damit wird ein erster wesentlicher Parameter benannt, in dem sich ein Institut

<sup>2</sup> Dies erfolgte in der Regel durch die Bildung von Anteilswerten für die Häufigkeit, mit der eine bestimmte Antwortkategorie einer Frage von den Abteilungen eines gegebenen Institutes gewählt wurde. Diese institutsspezifischen Anteilswerte gehen dann in Auszählungen oder Prozentuierungen über alle Institute hinweg mit einem Gewichtungsfaktor ein, welcher dem Anteil der Antwortkategorie über alle Abteilungen des jeweiligen Institutes entspricht. Wenn z.B. nur zwei von drei Abteilungen des Institutes X bei einer Frage die zweite Antwortkategorie wählten, die dritte Abteilung aber die erste Kategorie angab, wird Institut X bei der Häufigkeitsauszählung für die zweite Antwortkategorie dieser Frage nicht mit einer Häufigkeit von ‚1‘ gewertet, sondern mit einer Häufigkeit von ‚0,6666‘. Entsprechend wird Institut X für die erste Antwortkategorie mit ‚0,3333‘ gezählt.

mit seinen Strategien zur Datenhandhabung bewegt, denn technische, rechtliche und praktische Gegebenheiten des Forschungsdatenmanagements unterscheiden sich stark nach Quelle, Beschaffenheit und Volumen der anfallenden Daten.

Weil Klassifikationen von Forschungsdatentypen nicht existieren, wurde diese Einleitungsfrage offen gestellt, d.h., die Befragten wurden zur Nennung von Stichworten aufgefordert, die erst im Nachgang kategorisiert wurden. Es wurden dabei im ersten Schritt aus den offenen Antworten zu **Frage 1** (mit bis zu drei Antwortoptionen) Fein-Kategorien für Forschungsdatentypen gebildet. Diese immer noch große Anzahl von Kategorien wurde der Übersichtlichkeit halber in einem zweiten Schritt zu Grobkategorien zusammengefasst. Analog wurde bei weiteren offenen Fragen vorgegangen. Auf diesen Grob-Kategorien beruhen die meisten hier berichteten Auswertungen.

F1. Forschungsdatentypen in der Leibniz-Gemeinschaft

Grobkategorie	Feinkategorie	Absolute Anzahl	Grobkategorie: Anteil an Instituten (%)
			Feinkategorie: Anteil in allen Abteilungen (%)*
1. Quantitative Mess-/Erhebungsdaten		50	64,9
	<i>quantitative Messdaten</i>		48,2
	<i>amtliche Statistiken</i>		8,7
	<i>Umfrageantw. (quant.)</i>		25,4
	<i>Zeitreihendaten</i>		4,8
	<i>Aggregatdaten</i>		3,0
2. Komplexe Beobachtungsdaten		23	29,9
	<i>Komplexe Analysedaten</i>		15,2
	<i>Geodaten</i>		8,7
	<i>Gendaten</i>		2,7
	<i>experiment. Daten</i>		9,6
3. Komplexe synthetische Daten		10	13,0
	<i>Algorithmen</i>		4,5
	<i>Modelle</i>		9,1
4. Qualitative Daten		17	22,1
	<i>qualitative Daten</i>		4,2
	<i>Umfrageantw. (qual.)</i>		5,7
	<i>Textdaten</i>		17,1
5. AV-Materialien		27	35,1
	<i>Audiodaten</i>		11,4
	<i>Bilder</i>		28,5
	<i>Filme</i>		12,5
6. Anderes		14	18,2
	<i>Metadaten</i>		2,6
	<i>sonstiges</i>		14,9

\* Wegen der Möglichkeit von Mehrfachnennungen können sich die Prozentwerte der Feinkategorien zu mehr als dem jeweiligen Prozentwert der zugehörigen Grobkategorie addieren sowie alle Prozentwerte zu mehr als 100 %.

Die folgende Tabelle zeigt die Verteilung der Grobkategorien nach Sektionen.

### F1. Forschungsdatentypen über die Sektionen

Sektion	Grobkategorie	Absolute Anzahl	Anteil an Instituten der Sektion (%)
A	1. Quantitative Mess-/Erhebungsdaten	6	42,9
	2. Komplexe Beobachtungsdaten	0	0
	3. Komplexe synthetische Daten	0	0
	4. Qualitative Daten	7	50,0
	5. AV-Materialien	10	71,4
	6. Anderes	4	28,6
B	1. Quantitative Mess-/Erhebungsdaten	16	94,1
	2. Komplexe Beobachtungsdaten	4	23,5
	3. Komplexe synthetische Daten	0	0
	4. Qualitative Daten	5	29,4
	5. AV-Materialien	2	11,8
	6. Anderes	1	5,9
C	1. Quantitative Mess-/Erhebungsdaten	11	61,1
	2. Komplexe Beobachtungsdaten	4	22,2
	3. Komplexe synthetische Daten	3	16,7
	4. Qualitative Daten	3	16,7
	5. AV-Materialien	6	33,3
	6. Anderes	3	16,7
D	1. Quantitative Mess-/Erhebungsdaten	10	55,6
	2. Komplexe Beobachtungsdaten	7	38,9
	3. Komplexe synthetische Daten	4	22,2
	4. Qualitative Daten	1	5,6
	5. AV-Materialien	7	38,9
	6. Anderes	4	22,2
E	1. Quantitative Mess-/Erhebungsdaten	7	70,0
	2. Komplexe Beobachtungsdaten	8	80,0
	3. Komplexe synthetische Daten	3	30,0
	4. Qualitative Daten	1	10
	5. AV-Materialien	2	20,0
	6. Anderes	2	20,0

\* Wegen der Möglichkeit von Mehrfachnennungen können sich die Prozentwerte zu mehr als 100 % addieren.

### F.2, Metadatenstandards

Eine zentrale Frage ist, ob in einer Disziplin überhaupt Metadatenstandards vorliegen bzw. hinreichend verbreitet sind. Wenn die Anbieter von Forschungsdaten sich bei der Erschließung auf standardisierte Erschließungsmethoden und Formate stützen können, ist eine sinnvolle und nutzerfreundliche Organisation der Datensätze in Verzeichnissen und Repositorien weit leichter

möglich. Nur 47 % der Befragten (gezählt auf Abteilungsebene, d.h. mit 136 Fällen) geben für ihre jeweilige Disziplin an, es gebe Metadatenstandards. Von diesen geben wiederum 58 % an, die existierenden Standards seien in ihrer Disziplin auch durchgesetzt. Ein in diesem Sinne gut vorbereitetes Arbeitsfeld finden somit nur 21 % der Datenanbieter vor.

Und selbst dies ist weiter einzuschränken. Bei der offenen Abfrage der Namen der in der jeweiligen Disziplin tatsächlich verwendeten Standards erweist sich, dass viele Befragte anstelle von Metadatenstandards alleine auf Dateiformate, Ontologien oder andere verwandte Konstrukte verweisen, die nicht die vollen Funktionen eines Metadatenstandards erfüllen.<sup>3</sup>

F2. Metadatenstandards in Disziplin nach Sektionen  
(Auswertung auf Institutsebene)

Sektion		N	Anteil (davon mit breiter Anwendung in der Disziplin)
A	F2. Hat Disziplin Metadatenstandards	13	42,3
	<i>F2a. Breite Anw. v. Metadatenstandard</i>	5	80,0
B	F2. Hat Disziplin Metadatenstandards	13	48,1
	<i>F2a. Breite Anw. v. Metadatenstandard</i>	8	75,0
C	F2. Hat Disziplin Metadatenstandards	16	50,1
	<i>F2a. Breite Anw. v. Metadatenstandard</i>	10	55,5
D	F2. Hat Disziplin Metadatenstandards	17	35,3
	<i>F2a. Breite Anw. v. Metadatenstandard</i>	6	58,3
E	F2. Hat Disziplin Metadatenstandards	10	80,0
	<i>F2a. Breite Anw. v. Metadatenstandard</i>	8	37,5

#### 4.2 Forschungsdaten in den Instituten

##### F.3, Erzeugung eigener Forschungsdaten

In neun der 77 teilnehmenden Institute werden in keiner Abteilung eigene Forschungsdaten erzeugt. Der Anteil der Institute, die Forschungsdaten erzeugen (wo verfügbar, gewichtet mit der Anzahl der internen Abteilungen, die dies tun), liegt damit insgesamt bei **88 %**. Damit ist die Erzeugung eigener Daten zwar deutlich der Normalfall, aber keineswegs eine Selbstverständlichkeit.

**Kommentar [MQ1]:** Nachfragen:  
welche Institute erzeugen keine Daten?  
Gibt es da Typen/Gründe?

<sup>3</sup> Dies demonstriert freilich auch die Schwierigkeit der Erfassung des Themas mit seinen wenig kanonisierten Begrifflichkeiten in einem standardisierten Fragebogen.

### F3. eigene Forschungsdaten

Sektion	N	Anteil
A	14	82,1
B	17	74,5
C	18	94,4
D	18	77,8
E	10	100,0

#### F.4, Umgang mit eigenen Forschungsdaten

Frage 4 a-d richtet den Blick nun auf die Regularien, welche die Institute ihren Mitarbeitern beim Umgang mit Forschungsdaten gegebenenfalls vorgeben. Die vorgegebenen Antwortoptionen lassen sich dabei grob nach zunehmender Strenge dieser Regularien anordnen. (Antworten aus der offenen Antwortoption 4e wurden wo möglich den Antwortkategorien a-d nachträglich zugeordnet).

#### F4. Vorgegebene Regeln (Alle Institute)

Frage	N	Anteil
F4a gute Praxis	64	97,0
F4b Backup	62	48,9
F4c interne Nachweisdatenbank	63	39,6
F4d öffentlicher Zugang	61	24,2

Wie nicht anders zu erwarten, ist die Einhaltung der guten wissenschaftlichen Praxis in nahezu allen Instituten eine Selbstverständlichkeit – auch bei den zwei Instituten, für die diese Antwortoption nicht gewählt wurde, dürfte wohl eher ein abweichendes Verständnis der Frageformulierung der Grund sein als die fehlende Bereitschaft der Mitarbeiter zur Befolgung der üblichen wissenschaftlichen Normen. Massive Differenzierungen zeigen sich jedoch bereits bei der zweiten Antwort: weniger als die Hälfte der Institute gibt ihren Mitarbeitern verbindlich physische Backups der Daten vor, bei knapp 40% der Institute werden die im Haus erzeugten Forschungsdaten systematisch erfasst, und nur ein knappes Viertel der Institute fordert von seinen Mitarbeitern, die Daten über Repositorien Dritten systematisch zugänglich zu machen.

#### F4. Regeln (nach Sektion)

Sektion (Frage)		N	Anteil
A	F4a gute Praxis	11	95,5
	F4b Backup	10	20,0
	F4c interne Nachweisdatenbank	11	27,3
	F4d öffentlicher Zugang	9	11,1
B	F4a gute Praxis	14	92,9
	F4b Backup	15	43,3
	F4c interne Nachweisdatenbank	15	53,3
	F4d öffentlicher Zugang	15	42,2
C	F4a gute Praxis	17	97,5
	F4b Backup	16	64,6
	F4c interne Nachweisdatenbank	17	37,8
	F4d öffentlicher Zugang	17	20,2
D	F4a gute Praxis	12	100,0
	F4b Backup	11	50,0
	F4c interne Nachweisdatenbank	10	25,0
	F4d öffentlicher Zugang	10	20,0
E	F4a gute Praxis	10	100,0
	F4b Backup	10	60,0
	F4c interne Nachweisdatenbank	10	50,0
	F4d öffentlicher Zugang	10	20,0

#### F.5, Zwecke der internen Regeln

Befragt nach dem Zweck der Verabschiedung von Regularien geben die Vertreter der Institute weit überwiegend mehrere Ziele an. Dominanter Zweck ist dabei die Re-Analyse der Daten im eigenen Institut, entweder durch die ursprünglichen oder andere Forscher. Drei Viertel der Institute wollen mit den Regeln aber auch die Replizierbarkeit von Analysen sichern, also einem methodischen Grundprinzip der Wissenschaft dienen. 69 % wollen rechtliche Risiken etwa aus dem Gebiet des Datenschutzes ausschließen, noch die Hälfte der Institute verfolgt das Ziel, die Daten als Kulturgut oder zeitspezifische bzw. historische Beschreibung zu bewahren.

Antwortoption	N	Anteil
F5a Replikation	62	75,1
F5b eig. Re-Analyse	66	96,2
F5c Re-Analyse im Institut	65	95,7
F5d langfristige Informationserhaltung	61	50,3



F5e Ausschluss Datenschutz- /Sicherheitsrisiken 60 68,8

*F5f es gibt keine Datenschutzrisiken* 24\* 45,1\*\*

\*Anzahl an Befragten, die bei F.5e *nicht* Datenschutz-/Rechtsrisiken anführen

\*\*Anteil an Befragten, die bei F.5e *nicht* Datenschutz-/Rechtsrisiken anführen und dies mit dem Fehlen solcher Risiken begründen

Sektion (Frage)		N	Anteil
A	F5a Zweck Replikation	10	45,0
	F5b Zweck eig. Re-Analyse	11	86,4
	F5c Zweck Re-Analyse im Inst.	11	100,0
	F5d Zweck Info.-Erhaltung	11	68,2
	F5e Ausschluss Risiken	10	80,0
	F5f keine Risiken	2*	50,0**
B	F5a Zweck Replikation	14	69,6
	F5b Zweck eig. Re-Analyse	15	100,0
	F5c Zweck Re-Analyse im Inst.	14	92,3
	F5d Zweck Info.-Erhaltung	14	35,7
	F5e Ausschluss Risiken	15	73,3
	F5f keine Risiken	5*	10,0**
C	F5a Zweck Replikation	16	89,6
	F5b Zweck eig. Re-Analyse	17	94,1
	F5c Zweck Re-Analyse im Inst.	17	89,9
	F5d Zweck Info.-Erhaltung	15	54,4
	F5e Ausschluss Risiken	13	71,4
	F5f keine Risiken	7*	47,6**
D	F5a Zweck Replikation	13	76,9
	F5b Zweck eig. Re-Analyse	13	100,0
	F5c Zweck Re-Analyse im Inst.	13	100,0
	F5d Zweck Info.-Erhaltung	11	36,4
	F5e Ausschluss Risiken	13	53,9
	F5f keine Risiken	6*	50,0**
E	F5a Zweck Replikation	9	88,9
	F5b Zweck eig. Re-Analyse	10	100,0
	F5c Zweck Re-Analyse im Inst.	10	100,0

F5d Zweck Info.-Erhaltung	10	60,0
F5e Ausschluss Risiken	9	66,7
F5f keine Risiken	4*	75,0**

\*Anzahl an Befragten, die bei F.5e *nicht* Datenschutz-/Rechtsrisiken anführen

\*\*Anteil an Befragten, die bei F.5e *nicht* Datenschutz-/Rechtsrisiken anführen und dies mit dem Fehlen solcher Risiken begründen

#### F.14, Policy zum Datenmanagement

33 % der Institute haben eine ausformulierte Policy zum Forschungsdatenmanagement in mindestens in einer ihrer Abteilungen, d.h. in 67 % der Institute wird nicht von einer solchen Policy berichtet. Gewichtet man dies mit den jeweiligen Abteilungsangaben, liegt der durchschnittliche Anteil von Instituten *mit* Datenmanagement-Policy bei 26 %.

Sektion	N	Anteil
A	14	10,7
B	17	33,3
C	18	14,3
D	18	36,1
E	10	40,0

#### 4.3 Forschungsdatenrepositories in der Disziplin

Die Fragen zu Angeboten in der Disziplin *jenseits* der jeweiligen Institute haben ausdrücklich nicht die Institute selbst zum Gegenstand. Daher wird hier auf die Aggregation über Abteilungen verzichtet und auf die Gesamtstichprobe von 136 Fällen zurückgegriffen.

**Kommentar [MQ2]:** In diesem Abschnitt ist ggf. noch die Disziplin группierung anzuwenden

#### F.11, Drittanbieter existieren

59 % der Befragten, das sind 80 Fälle, berichten, dass in der Disziplin, aber außerhalb ihres eigenen Instituts/ihrer eigenen Abteilung ein Forschungsdatenrepository gibt. Solche Repositorien sind am meisten bekannt in den Instituten der Sektionen B und E, an wenigsten bekannt in den Instituten der Sektionen C und D.

##### F.11, externes Repository vorhanden

Sektion	N	Anteil
A	19	68,4
B	31	77,4
C	54	46,3
D	20	45,0
E	12	75,0

### F.12, Details zum wichtigsten Dritt-Repository

Da Frage 12 mit den Unterfragen a-f nur an 72 Personen gestellt werden konnte und zu den Detailfragen viele der Befragten keine gültigen Antworten geben konnten, ist eine Untergliederung der Antworten nach Sektionen nicht mehr sinnvoll. Die folgende Tabelle zeigt jeweils den Anteil einer der beiden inhaltlichen Antwortkategorien zu den Unterfragen F.12a-f.

F.12 a-f, Details zu externem Repository  
(Auswertung auf Befragtebene)

Frage	N (gültige Antworten)	Anteil ‚ja‘ bzw. ‚national‘ (vs. international)
F.12a Nutzung durch Mitarbeiter des eig. Instituts	72	73,6
F.12b Nutzung mit relevanten Kosten	72	8,3
F.12c Daten im Rep. sind vorwiegend national (vs. International)	68	38,2
F.12d Rep. hat differenzierte Zugriffsrechte für Nutzerkreise	52	50,0
F.12e Rep. hat Embargo-Regeln	41	46,3
F.12f Standards im Rep. national (vs. international)	46	4,3

### 4.4 Daten-Repositories der Institute

#### F.6, Anzahl von Repositories

Insgesamt betreiben 33,1 % der Abteilungen (45) bzw. 37,7 % (29) der Institute 89 Daten-Repositories unterschiedlicher Art. Entsprechend gibt es 48 Institute, die kein Repository betreiben.

F.6a, Anzahl der Repositories

Sektion	N (Anzahl Institute)	Anzahl Institute mit	
		Rep.	Anzahl Repositories
A	14	6	15
B	17	5	9
C	18	8	36
D	18	6	14
E	10	4	15
Gesamt	77	23	89

21 weitere Institute bzw. Abteilungen geben auf Frage 6b an, dass die Einrichtung eines Repository (für interne oder externe Nutzer) in ihrer Abteilung/ihrem Institut bereits fest geplant sei. In der planbaren Zukunft werden die Institute der Leibniz-Gemeinschaft also über 100 unterschiedliche Datenrepositories betreiben.

#### F.7, Details Repositories

Die Natur der bereits bestehenden Instituts-Repositories ist erwartungsgemäß gemischt. Überwiegend, aber keineswegs durchgängig, werden Metadatenstandards beachtet und die Daten durch Verschlagwortung für eine offene Suche erschlossen. Die Verwendung von sog. Persistenten Identifikatoren und die damit immer im Zusammenhang zu sehende Versionierung von

Forschungsdatensätzen ist in knapp 44 % der Institute bzw. Abteilungen mit Repository in mindestens einem der Repositories des Instituts Praxis, jedoch enthielten die offenen Antworten zu dieser Frage auch knapp 20 % Zweifelsfälle.

F.7 c-f, Details zu Repositories  
(Auswertung mit Bezug auf Repositories in Abteilungen; bis auf F.7f)

Frage	N (gültige Antworten)	
F.7c konform mit Metadatenstandard	45	Anteil ‚ja‘ an allen Repositories: 62,9%
F.7d Daten systematisch durch Schlagworte beschrieben	45	Anteil ‚ja‘ an allen Repositories: 61,8%
F.7e Anzahl von Einträgen (Datensätzen)	35	Range: 3 - 16.000.000, Mittelwert: 1.222.975
F.7f persistente Identifikation/Versionierung (nur einmal für alle Repositories eines Institutes kodiert!)	46	Anteil ‚ja‘ an allen Instituten: 43,5%
		Anteil ‚unklar‘ an allen Instituten: 19,6%

**Kommentar [MQ3]:** Hier liegt ein Problem in der Datenkodierung der offenen Antworten vor: die Information ist im Datensatz nicht für jedes einzelne Repository aufgelöst, obwohl die Information im Fragebogen so vorliegt. Bei Bedarf kann dies nachkodiert werden.

#### F.8, Verfügbarkeit von Repositories für Externe; F.9, Details von Repositories für Externe

Insgesamt 30 der 89 Repositories (34 %) sind für die Nutzung durch Instituts-Externe Forscher offen. Von diesen nehmen 23 Repositories (25,8 % von 89) auch Daten auf, die von externen Forschern erzeugt wurden, sie haben also eine „Verteilerfunktion“ für Forschungsdaten für die Gesamt-Community. Bei 14 Repositories ist dies nach den Antworten zu Frage 9.e sogar der Hauptzweck.

#### F.10, Freigabe von Repositories für Externe frei, gebührenpflichtig, rechtlich notwendig

Zu 44 Repositories wird berichtet, die Bereitstellung der Forschungsdaten ohne Gebühren sei das Hauptziel. Nur einmal wird angeführt, dass ein finanzieller Deckungsbeitrag durch Gebühren auf den Datenzugriff ein Ziel sei, für das ein Repository betrieben wird.

#### 4.5 Andere Dienste der Institute

Knapp 39 % der Institute (30) bieten (in mindestens einer Abteilung) jenseits von Repositories andere oder weitere Dienste zur Forschungsdatenhandhabung an. Die speziellen Dienstleistungen dabei werden in der folgenden Tabelle (Anteile von Diensten gewichtet durch Instituts-interne Verteilung über Abteilungen) in der Übersicht dargestellt.

F.13, a-d weitere Dienste		
	N	Anteil von Diensten
F13 weitere Dienste	77	39,0
F13a Dienst pers. ID für eig. Daten	30	13,3
F13b Dienst pers. ID für externe Daten	30	16,7
F13c Nachweis Forschg.daten im Bib.-System	30	26,7
F13d Tools für Ablage	30	10,0

F.13, a-d weitere Dienste nach Sektionen

Sektion		N	Anteil
A	F13 weitere Dienste	14	53,6
	F13a Dienst pers. ID intern	8	12,5
	F13b Dienst pers. ID extern	8	12,5
	F13c Nachweis For.Metadaten im Bib.-System	8	31,3
	F13d Tools für Ablage	8	0,0
B	F13 weitere Dienste	17	21,6
	F13a Dienst pers. ID intern	5	20,0
	F13b Dienst pers. ID extern	5	20,0
	F13c Nachweis For.Metadaten im Bib.-System	5	40,0
	F13d Tools für Ablage	5	0,0
C	F13 weitere Dienste	18	35,5
	F13a Dienst pers. ID intern	9	3,7
	F13b Dienst pers. ID extern	9	14,8
	F13c Nachweis For.Metadaten im Bib.-System	9	11,1
	F13d Tools für Ablage	9	13,3
D	F13 weitere Dienste	18	25,0
	F13a Dienst pers. ID intern	5	20,0
	F13b Dienst pers. ID extern	5	0,0
	F13c Nachweis For.Metadaten im Bib.-System	5	20,0
	F13d Tools für Ablage	5	20,0
E	F13 weitere Dienste	10	30,0
	F13a Dienst pers. ID intern	3	0,0
	F13b Dienst pers. ID extern	3	0,0
	F13c Nachweis For.Metadaten im Bib.-System	3	33,3
	F13d Tools für Ablage	3	0,0

## **5 Schlussfolgerungen**

### **5.1 *Die Situation***

### **5.2 *Handlungsempfehlungen***

## **6 Anhänge**