

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

Bedarfsplanung eines institutionellen Repositoriums für Forschungsdaten

Robert Strötgenⁱ

2019

Zitiervorschlag

Strötgen, Robert. 2019. Bedarfsplanung eines institutionellen Repositoriums für Forschungsdaten. *Bausteine Forschungsdatenmanagement. Empfehlungen und Erfahrungsberichte für die Praxis von Forschungsdatenmanagerinnen und -managern* Nr. 2/2019: S. 112-120. DOI: [10.17192/bfdm.2019.2.8106](https://doi.org/10.17192/bfdm.2019.2.8106).

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

ⁱTechnische Universität Braunschweig. ORCID: [0000-0003-3320-5187](https://orcid.org/0000-0003-3320-5187)

1 Zusammenfassung

Das Forschungsdaten-Team der TU Braunschweig (bestehend aus Universitätsbibliothek, Rechenzentrum und Forschungsservice) hat im Januar/Februar 2018 eine Online-Befragung zum Thema Forschungsdatenmanagement durchgeführt. Anlass war die Bedarfsplanung für die Einrichtung eines institutionellen Forschungsdatenrepositoriums. Im Detail wurde der Bedarf für die Archivierung und Publikation von Forschungsdaten abgefragt. Bei dieser Gelegenheit wurde ebenfalls der Bedarf auf weiteren Unterstützungsgebieten (Beratung, Datenmanagementpläne, rechtliche Fragen etc.) einbezogen.

Dieser Beitrag stellt den Kontext und die Ergebnisse der Bedarfsabfrage dar. Dabei wird besonders auf die Einschränkungen bei der Erhebung und Auswertung eingegangen. Besonderes Gewicht haben Effekte wie die Sichtbarkeit und die Diskussion des Themas.

Dieser Beitrag basiert auf einem Vortrag im Rahmen des DINI/nestor-Workshops "Bedarfserhebungen - Grundlage für passgenaue Infrastrukturen" am 29./30. Januar 2019.

2 Ausgangslage

Die Technische Universität Braunschweig hat im Herbst 2016 mit dem Aufbau von Service- und Unterstützungsstrukturen für das Forschungsdatenmanagement begonnen, zunächst am Bedarf orientiert mit Beratungen vor allem im Kontext von Drittmittelanträgen. Im Januar 2017 wurde die Kooperation der Forschungsdatenverantwortlichen der TU9¹ begonnen. Mittlerweile wurden aus dieser Zusammenarbeit heraus z.B. Weiterbildungsmaterialien gemeinsam entwickelt und zur Nachnutzung publiziert.² Seit Frühjahr 2017 existiert an der TU Braunschweig ein gemeinsames Team für Forschungsdaten unter Federführung der Universitätsbibliothek mit Beteiligung des Rechenzentrums (GITZ), des Forschungsservices und des CIO.³ Im Dezember 2017 wurde im Senat der TU Braunschweig eine Leitlinie zum Umgang mit Forschungsdaten verabschiedet, die vorher in allen Fakultäten und weiteren Gremien ausführlich diskutiert wurde.⁴ Seit Februar 2018 führt die UB Braunschweig regelmäßig Weiterbildungen zum Forschungsdatenmanagement durch.

¹<https://www.tu9.de/>, Zugriff am 28.10.2019 – TU9 ist die Allianz führender Technischer Universitäten in Deutschland.

²Vergleiche die Beiträge unter "Community for RDM Training Materials for the Engineering Sciences" in Zenodo, Zugriff am 28.10.2019, https://zenodo.org/communities/rdm_training_engineering_sciences/.

³S. a. die Webseite <https://www.tu-braunschweig.de/forschung/daten/forschungsdaten/>, Zugriff am 28.10.2019.

⁴"Leitlinie zum Umgang mit Forschungsdaten an der Technischen Universität Braunschweig", Technische Universität Braunschweig, Zugriff am 28.10.2019, <https://ub.tu-braunschweig.de/publizieren-openaccess/forschungsdaten/forschungsdatenleitlinie.pdf>.

In der Leitlinie zum Forschungsdatenmanagement sagt die TU Braunschweig den Forschenden Unterstützung u.a. durch Infrastrukturangebote zu. Obwohl fachliche Forschungsdatenrepositorien in aller Regel für die Publikation von Forschungsdaten empfohlen werden, gibt es immer wieder Gründe, warum ein institutionelles Repositorium benötigt wird (z.B. für die reine Archivierung von Forschungsdaten oder bei Projekten mit Industriepartnern, die keine Publikation der Forschungsdaten ermöglichen). Deshalb ist die Bereitstellung eines institutionellen Forschungsdatenrepositoriums ein wichtiges Projekt, das 2016 zur Förderung im Präsidium der TU Braunschweig beantragt wurde.

3 Vorbereitung und Durchführung der Bedarfsabfrage

Die AG-Forschung des IT-Lenkungsausschusses der TU Braunschweig unter der Leitung des Vizepräsidenten für Forschung steuert die Aktivitäten zum Forschungsdatenmanagement. Aus dem Kreis der AG kam die Anregung, in der Vorbereitung des Projekts den Bedarf der Forschenden abzufragen. Im kollegialen Austausch unter anderem im Kreis der TU9 war die Schwierigkeit von Forschenden, den Bedarf im Voraus abzuschätzen, bereits mehrfach diskutiert worden, so dass eine erhebliche Skepsis bestand, durch eine solche Bedarfsabfrage belastbare Planungsdaten zu erhalten. Trotzdem wurde der Vorschlag aus verschiedenen Gründen angenommen:

- Eine Bedarfsabfrage unter allen Forschenden kann Legitimation für das Projekt herstellen bzw. dokumentieren.
- Eine Bedarfsabfrage kann bisher unbekannte Anforderungen sichtbar machen.
- Eine Bedarfsabfrage unter allen Forschenden der TU Braunschweig erhöht die Sichtbarkeit des Themas sowie der Ansprechpartner und kann die Awareness für das Forschungsdatenmanagement erhöhen.
- Eine Bedarfsabfrage kann ein Anlass für Forschende sein, den eigenen Umgang mit Forschungsdaten zu reflektieren.

Deshalb wurde im Herbst 2017 eine Abfrage zum Bedarf bei Archivierung und Publikation von Forschungsdaten sowie weiteren Dienstleistungen zum Forschungsdatenmanagement konzipiert und auf der Basis des vom Rechenzentrum bereitgestellten Werkzeugs LimeSurvey als Online-Umfrage umgesetzt. Die Abfrage wurde mit dem Datenschutzbeauftragten abgestimmt, der auf eine anonyme Antwortmöglichkeit bestand. Die Angabe von Kontaktdaten war freiwillig möglich. Die Umfrage wurde über mehrere Kanäle (u.a. Mail-Verteiler der Fakultäten und Forschungszentren, Blog-Beiträge, Beitrag im Magazin der TU Braunschweig⁵) beworben. Nach ungefähr der Hälfte der Laufzeit wurde eine zweite Werbekation zur Erinnerung gestartet. Die Befragung lief vom 18.01. bis 23.02.2018. Insgesamt wurden bei ca. 3.700 hauptberuflichen

⁵<https://magazin.tu-braunschweig.de/magazin/>, Zugriff am 28.10.2019.

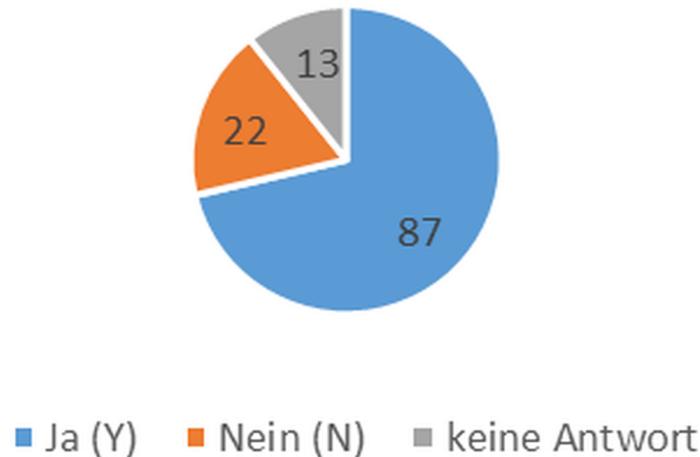


Abbildung 1: Archivierung Forschungsdaten (absolute Zahlen)

Mitarbeiterinnen und Mitarbeitern 240 Fragebögen abgesendet, von diesen waren 85 vollständig (also kein Abbruch vor der letzten Frage).

Die Bedarfsabfrage enthielt drei inhaltliche Blöcke:

- Archivierung von Forschungsdaten
- Publikation von Forschungsdaten
- Weitere Services im Kontext Forschungsdaten

Dabei wurden nicht nur (quantitative) Fragen zu Speicherbedarf etc. gestellt, sondern auch nach bereits genutzten Diensten und bekannten Infrastrukturen gefragt.

4 Ausgewählte Ergebnisse und Interpretation

Im Ergebnis zeigte sich u.a., dass nahezu drei Viertel der Teilnehmenden in den nächsten drei Jahren Forschungsdaten erwarten, die archiviert werden sollen oder müssen (siehe Abbildung 1).

Mehr als zwei Drittel äußerten einen Bedarf an einem institutionellen Forschungsdatenrepositorium (der TU Braunschweig) für die Archivierung (siehe Abbildung 2).

Besonders hoch war mit über 80% der Anteil der Befragten, die in den nächsten drei Jahren eine Forschungsdatenpublikation erwarten (siehe Abbildung 3).

Bemerkenswert waren auch die Antworten auf die Frage, wo Forschungsdaten publiziert werden sollten. Hier erhielten fachliche Repositorien, Verlags-Repositorien und

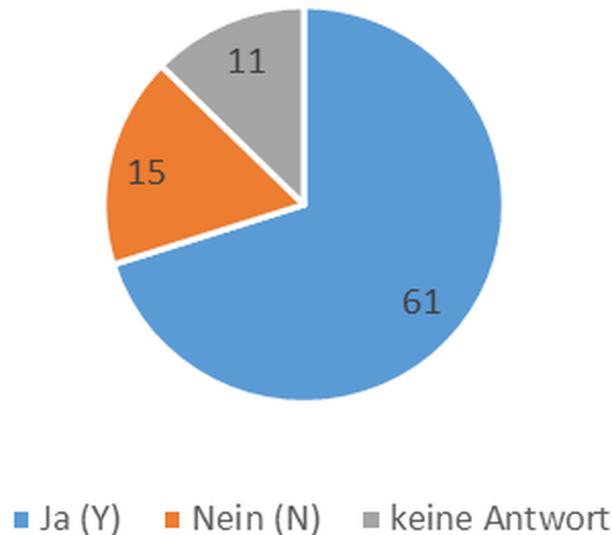


Abbildung 2: Bedarf institutionelles Forschungsdatenrepositorium (absolute Zahlen)

ein institutionelles Repositorien nahezu exakt gleich hohe Werte, lediglich die Option "institutionelles Repositorium einer Partnereinrichtung" erhielt nur wenige Antworten.

Ungefähr ein Drittel der Befragten gab an, fachliche Forschungsdatenrepositorien zu kennen, und ein knappes Viertel hat ein solches fachliches Repositorium bereits genutzt.

Für den Anlass der Abfrage war das erwartete Volumen an Forschungsdaten natürlich besonders relevant. Für die nächsten drei Jahre wurde bei der Archivierung ein Bedarf an ca. 200 TByte geäußert, bei der Publikation - wenig überraschend - mit ca. 10 TByte deutlich weniger. Hier wurde allerdings eine deutliche Verzerrung durch die anonyme Antwortmöglichkeit deutlich. Betrachtet man nur die nicht anonymen Antworten, beträgt der Archivierungsbedarf lediglich gut 80 TByte. Für die Publikation ist die Abweichung erheblich geringer. Da die Archivierung zu durchaus erheblichen Kosten führt, bleibt es bei allen anderen, bereits genannten Einschränkungen besonders problematisch, diese Kosten auf der Grundlage anonymer Äußerungen zu kalkulieren.

Versucht man (deutlich vereinfachend) den Bedarf linear weiter zu kalkulieren, wird in der Kumulation (zumindest in den ersten 10 Jahren werden keine Daten gelöscht) der Umfang der Daten zu erheblichen Kosten führen (siehe Abbildung 4/5):

Diese Kalkulationen sind natürlich sowohl in Bezug auf das Wachstums des Volumens an Forschungsdaten als auch bezüglich der unveränderten Kosten sehr vereinfachend. Trotzdem sind sie mit Hinweis auf alle Einschränkungen in die weitere Planung eingeflossen. Für die Auswertung und Belastbarkeit der Umfrage ergeben sich

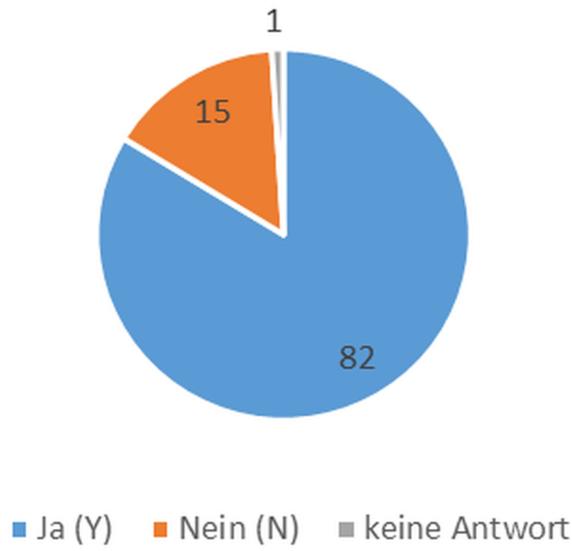


Abbildung 3: Publikation Forschungsdaten (absolute Zahlen)

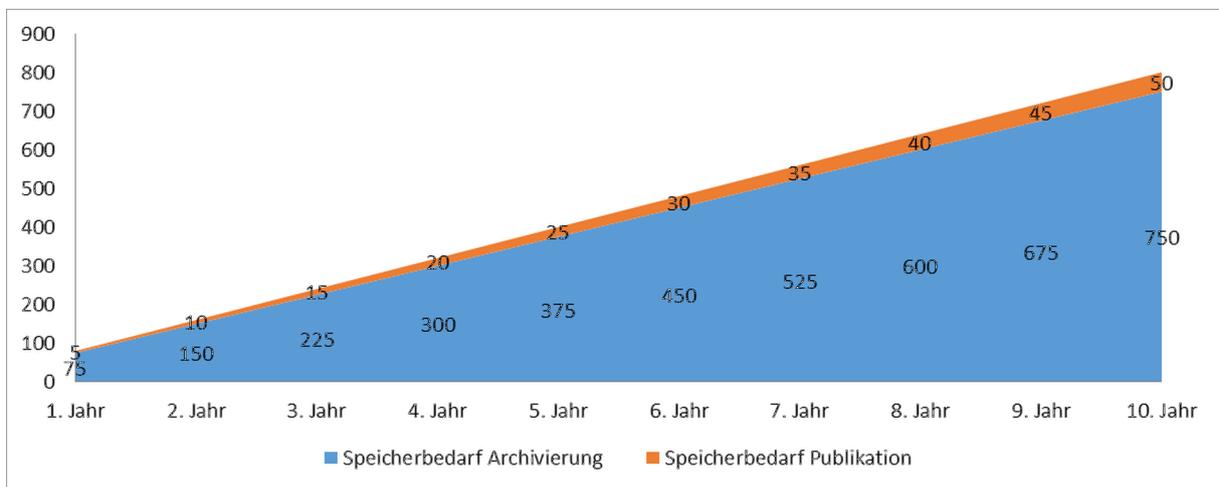


Abbildung 4: Kalkulation Speicherbedarf (auf der Basis von 35 Nennungen)

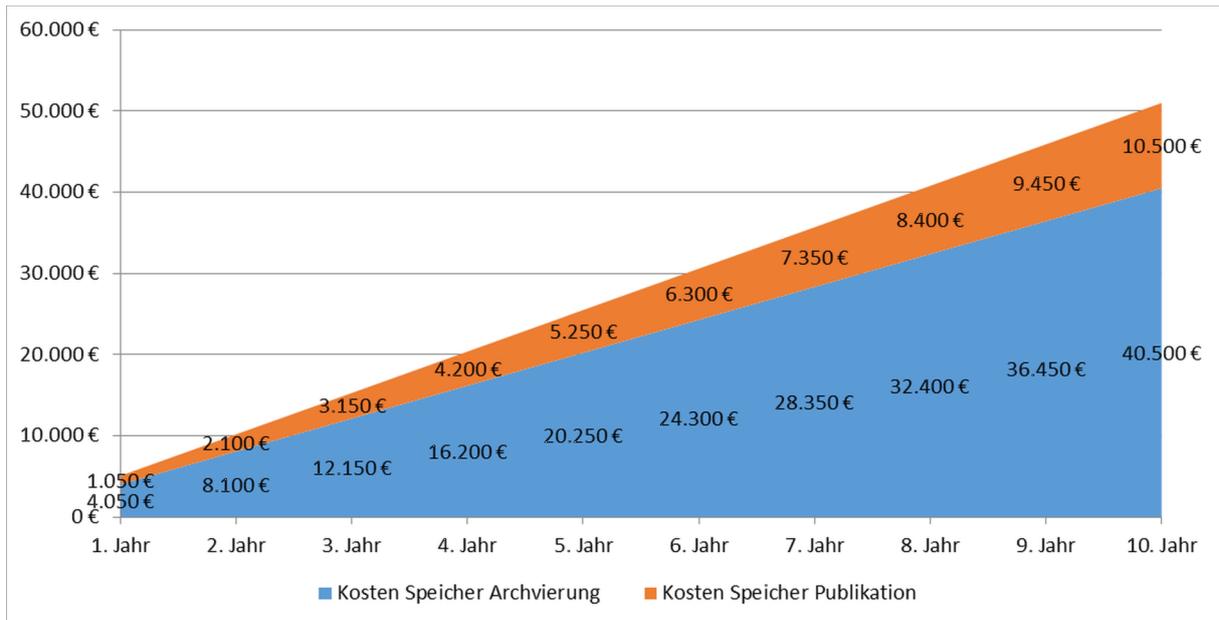


Abbildung 5: Kalkulation Speicherkosten (auf der Basis von 35 Nennungen)

weitere Einschränkungen:

- Durch die Möglichkeit, die Umfrage abubrechen, ergeben sich nicht konsistente Zahlen.
- Die Zahl der angesprochenen Forschenden ist nicht genau definiert (wissenschaftliche Mitarbeiter und Mitarbeiterinnen, aber auch Doktoranden und Doktorandinnen etc.). Die Rücklaufquote ist daher schwer einzuschätzen. Auf eine Hochrechnung des Volumens von den Antworten auf alle Forschenden wurde daher verzichtet.
- Die Motivation zur Antwort war sicherlich bei denen am höchsten, die ein Interesse am Thema Forschungsdaten haben. Insofern dürfte die Auswahl der Antwortenden nicht repräsentativ sein.
- Es ist aber auch nicht zu erwarten, dass alle Forschenden mit Unterstützungsbedarf beim Forschungsdatenmanagement geantwortet haben.

Die Umfrage wurde auch genutzt, um den Beratungs- und Unterstützungsbedarf nach weiteren Themen des Forschungsdatenmanagements abzufragen. Hier zeigte sich ein breit verteilter Bedarf über verschiedene Themen. Besonders häufig wurden rechtliche Fragen, Datenstandards, Datenmanagementpläne und Schulungen genannt (siehe Abbildung 6)⁶.

Die Umfrage erlaubte es auch, in offenen Antworten weitere Themen zu benennen. Neben allgemein wissenschaftspolitischen Themen ("Wer entscheidet denn, ob Forschungsdaten publiziert werden sollen oder müssen?" ; "Werden den Forschenden

⁶Sechster Balken: „Suche nach einem geeigneten Forschungsdatenrepositorium“.

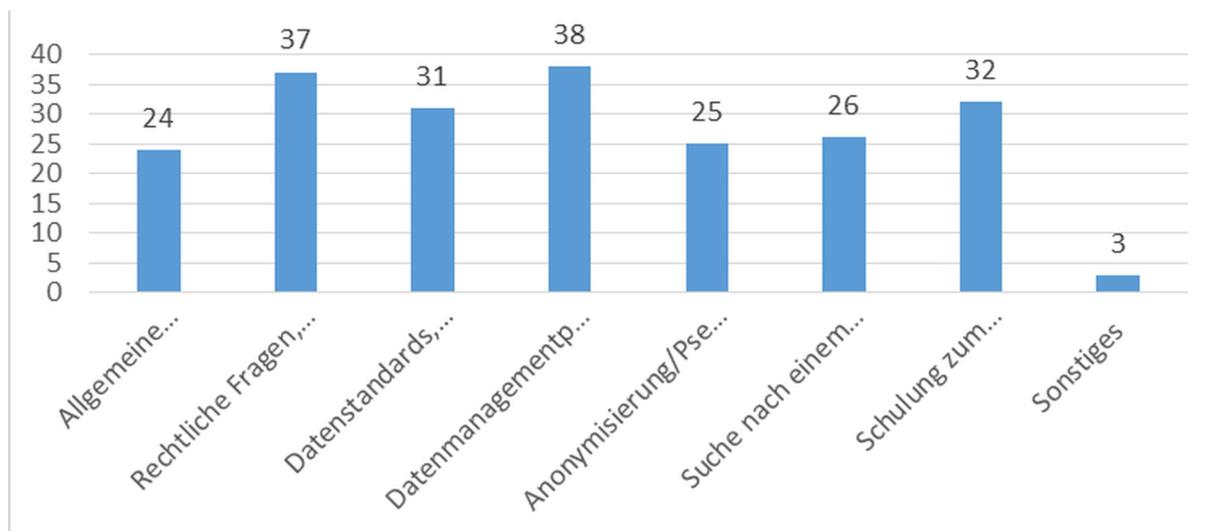


Abbildung 6: Unterstützungsbedarf (absolute Zahlen)

durch das Forschungsdatenmanagement Ressourcen entzogen und zusätzliche Pflichten aufgebürdet?“) spielten besonders Datenschutzfragen (z.B. zur Anonymisierung) eine Rolle.

5 Fazit

Im Fazit lässt sich zusammenfassen, dass die Bedarfsabfrage trotz aller Bedenken erfolgreich war.

Die Befragung war vom Aufwand her gut zu bewältigen und mit dem Online-Werkzeug LimeSurvey gut umsetzbar.

Das Ergebnis der Befragung hat dem internen Projekt zur Implementierung eines institutionellen Forschungsdatenmanagements zusätzliches Gewicht verliehen und Bedenken ausgeräumt. Das Thema Forschungsdatenmanagement wurde breit diskutiert und im Austausch mit Forschenden reflektiert. Die Beratungsdienstleistungen wurden bekannter und in der Folge häufiger (nahezu um 50% mehr) genutzt als vorher. Dem Projekt der Implementierung eines institutionellen Forschungsdatenrepositoriums wurde zusätzliche Legitimation gegeben.

Die Umfrage hat auch gezeigt, dass Forschungsdatenmanagement an der TU Braunschweig bereits stattfindet und ein erheblicher Bedarf an Infrastruktur und Services besteht. Für die Forschenden steht ein institutionelles Repositorium nahezu gleich bedeutend neben fachlichen Repositorien und Verlagsangeboten. Bisher unbekannte Bedarfe wurden nicht identifiziert, aber die Gewichtung hilft bei der Planung zusätzlicher Services.

Andere oder übergreifende Befragungen, die im Vorfeld ausgewertet wurden, ersetzen nicht die Erhebung in der eigenen Einrichtung in Bezug auf Legitimation und Diskussion des Themas.

Die Entscheidung des Präsidiums über die Einrichtung eines institutionellen Forschungsdatenrepositoriums, zu dessen Vorbereitung diese Befragung durchgeführt wurde, fiel im Dezember 2018 positiv aus.