

Anforderungsanalyse für ein universitätsweites nutzerorientiertes Forschungsdatenmanagementsystem basierend auf einer Nutzerumfrage

Sabine Dient, FernUniversität in Hagen, sabine.dient@studium.fernuni-hagen.de
Tanja Auge, Universität Regensburg, tanja.auge@informatik.uni-regensburg.de

Abstract Eine rasant ansteigende Datenmenge erfordert ein adäquates *Forschungsdatenmanagement* (FDM), welches den gesamten Forschungsdatenlebenszyklus abgedeckt. Dieser beinhaltet insbesondere das Sammeln, Beschreiben und Bereitstellen von Forschungsdaten. Ein *Forschungsdatenmanagementsystem* (FDMS) stellt die technische Umgebung dafür dar.

In der vorliegenden Studie haben wir die Bedürfnisse von potentiellen Nutzern eines universitätsweiten nutzerorientierten FDMS evaluiert. Mittels einer Literaturrecherche und einer Onlineumfrage ($n = 269$) wurden mögliche FDMS-Funktionen in Übereinstimmung mit den *FAIR-Prinzipien* bestimmt.

Einführung Forschungsdaten sollen FAIR [4] sein, d.h. **F**indable, **A**ccessible, **I**nteroperable, **R**eusable. Hieraus lassen sich bereits Anforderungen an ein FDMS schließen, das den Forschungsdatenlebenszyklus [1] (Abbildung 1) abbildet. Dieser geht von der Planung und der Sammlung der Daten, über die Verarbeitung und Analyse dieser, den Datenzugriff und die Archivierung bis zur Wiederverwendung der Daten. Zum Einen haben die einzelnen wissenschaftlichen Fachbereiche unterschiedliche Methoden zur Datenerhebung sowie unterschiedliche Verarbeitungsarten, woraus sich auch individuelle Bedürfnisse an ein FDMS ergeben. Zum Anderen wächst der Bedarf für ein FDMS durch neue Regelungen (z.B. eine verpflichtende Beschreibung eines FDM bei Förderanträgen an die DFG).

Ziel dieser Studie ist die Ermittlung von Nutzerbedürfnissen sowie die Ausarbeitung von Best Practices für ein universitätsweites nutzerorientiertes FDMS. Hierfür untersuchen wir insbesondere drei Hypothesen:

- (1) Die Präferenz für reichhaltige Features oder ein einfaches System ist abhängig von der technischen Ausrichtung des Fachgebiets.
- (2) Die Bereitschaft zur Verwendung eines FDMS ist abhängig von der Erfahrung mit FDM(S).
- (3) Ein FDMS wird von Digital Natives [3] bevorzugt.

Damit soll die Lücke geschlossen werden, die derzeit aufgrund einer Konzentration von Studien auf einzelne Fachbereiche statt einem fächerübergreifenden System besteht.

Design Die Studie ist als quantitative Analyse konzipiert, die in Form einer Onlineumfrage

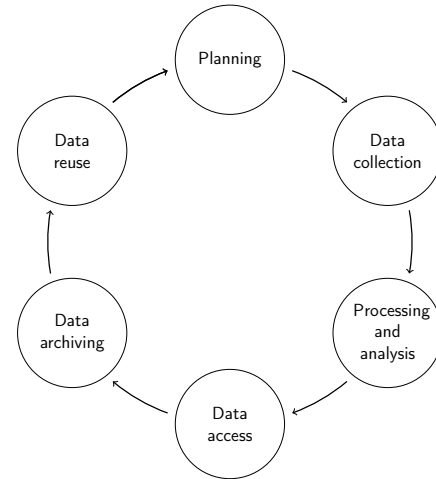


Abbildung 1: Forschungsdatenlebenszyklus

durchgeführt wurde. Über einen Zeitraum von acht Wochen sind 409 Fragebögen eingegangen, wovon 269 gültige Fragebögen – (vollständig) in angemessener Zeit ausgefüllt – waren. Diese kamen aus 37 Ländern (40 % Deutschland, 18 % Österreich, 9 % USA, 33 % andere Nationen). Die meisten Rückmeldungen stammen aus den Fachbereichen Naturwissenschaften (29,0 %), Sozialwissenschaften (27,1 %) sowie den technischen Wissenschaften (14,5 %). Die Humanmedizin und Gesundheitswissenschaften, die Agrarwissenschaften und Veterinärmedizin sowie die Geisteswissenschaften haben jeweils einen Anteil von rund 10 %.

Analyse der Nutzerbedürfnisse Basierend auf einer Literaturrecherche haben wir zunächst Kriterien an ein FDMS definiert, die wir später um weitere eigene Kriterien ergänzt haben. Diese konnten die Befragten auf einer Skala von 1 (gar nicht wichtig) bis 5 (sehr wichtig) bewerten. Die daraus resultierenden Top 3-Merkmale eines FDMS, basierend auf den Durchschnittswerten der gesamten Daten, sind:

- (1) Volle Kontrolle über eigene Daten (z.B. Löschung oder Export der Daten) mit 4,5/5
- (2) Verfolgung von Inhaltsänderungen mit 4,2/5
- (3) Datenexport in verschiedene Formate, Suchfilter (Facettensuche) und eindeutige Indikatoren für Daten mit 4,1/5.

Diese Top 3-Durchschnittswerte stimmen in weiten Teilen mit denen der oben genannten sechs Fachgebiete überein, wobei sich hauptsächlich die Reihenfolge unterscheidet.

Desweiteren haben wir die Teilnehmenden auch zu ihrer bevorzugten Art der Anwendung befragt. Zur Auswahl standen eine Webapplikation, eine mobile Applikation, eine lokale Installation oder eine beliebige Kombination daraus. Hier zeigten sich große Unterschiede bei den einzelnen Fachgebieten: Die Kombination einer Webapplikation mit einer lokalen Installation oder eine Webapplikation allein wurde von knapp 60 % der Befragten aller Fachgebiete gewünscht. Nur bei den Teilnehmenden der Agrarwissenschaften und der Veterinärmedizin lag dieser Anteil bei rund 25 %. 30 % dieser Gruppe präferieren ausschließlich eine lokale Installation. Der Ruf nach einer mobilen Applikation ist dagegen am größten bei den Befragten der Humanmedizin und der Gesundheitswissenschaften (14 %).

Überprüfung der Hypothesen Für die Überprüfung der drei aufgestellten Hypothesen haben wir statistische Methoden verwendet. Ziel war wiederum die Ableitung von Empfehlungen für die Erstellung eines FDMS.

Hypothese 1: Gegeben war eine Skala von 0 (einfaches System) bis 100 (System mit reichhaltigen Features). Obwohl die eher technisch orientierten Fachgebiete einen Durchschnittswert von 54 im Gegensatz zu 45 bei den weniger technischen Fachgebieten aufweisen, haben wir keinen Zusammenhang zwischen der technischen Ausrichtung und der Präferenz eines Fachgebiets nachweisen können. Somit ist die Hypothese, dass die Präferenz für reichhaltige Features oder ein einfaches System abhängig von der technischen Ausrichtung des Fachgebiets ist, widerlegt.

Hypothese 2: Von 106 Befragten, die vertraut mit FDM(S) sind, möchten 55 % ein FDMS für den Forschungsdatenlebenszyklus (Abbildung 1) verwenden. Unter den 97 Befragten, die entweder nicht damit vertraut oder sich nicht sicher sind, sind es nur 45 %. Somit ist die Hypothese, dass die Bereitschaft zur Verwendung eines FDMS abhängig von der Erfahrung mit FDM(S) ist, bestätigt.

Hypothese 3: Die Gruppe der unter 40-Jährigen (56 % der Befragten) möchte ein FDMS zu 61,8 % für die Phasen des Forschungsdatenlebenszyklus [1] verwenden. Bei Personen, die 40 Jahre und älter sind (44 %), liegt die Bereitschaft bei 67,5 %. Somit ist die Hypothese, dass ein FDMS von Digital Natives [3] bevorzugt wird, widerlegt.

Ergebnisse

- (1) Die Befragten zeigen eine mangelnde Vertrautheit mit FDM und FDMS.
- (2) Die Teilnehmenden wünschen sich eine Unterstützung für den gesamten Forschungsdatenlebenszyklus [1] (Abbildung 1), die Integration mit bestehender Software und benutzerfreundliche Oberflächen.

- (3) Digital Natives [3] zeigen keine höhere Bereitschaft für die Verwendung eines FDMS.
- (4) Die meisten wissenschaftlichen Fachbereiche bevorzugen eine lokale Installation in Kombination mit einer Webapplikation.
- (5) Die Nutzenden wünschen sich die Integration von häufig verwendeter Forschungssoftware.

Empfehlungen Die Studie zeigt, dass einer der wichtigsten Aspekte für ein universitätsweites nutzerorientiertes System eine Benutzeroberfläche ist, die abhängig von der wissenschaftlichen Disziplin der Nutzenden angepasst wird. Das System sollte zunächst lokal bereitgestellt werden und eine Integration mit beispielsweise Excel und SPSS (beliebteste Forschungssoftware) aufweisen. Zudem sollten die oben genannten Top 3-Merkmale eines FDMS, die auch die FAIR-Prinzipien erfüllen, bei der Entwicklung eines entsprechenden FDMS berücksichtigt werden.

Weitere Schlussfolgerungen Der niedrige Grad an Vertrautheit mit FDM (etwa 50 % der Befragten) zeigt einen Bedarf an Aus- und Fortbildung auf diesem Gebiet. Die Digital Natives könnten beispielsweise direkt durch eine Integration von FDM in das universitäre Curriculum unterstützt werden.

Für die Verbesserung der Auffindbarkeit der Daten und die Erstellung/Verarbeitung von Metadaten können bei der Implementierung Ontologien beachtet werden. Für den Entwurf eines FDMS basierend auf den Ergebnissen der Interviewstudie, insbesondere der ermittelten FDMS-Merkmale, empfehlen wir einen horizontalen Prototypen.

Danksagung Dieser Kurzbericht basiert auf einer gleichnamigen Bachelorarbeit [2], die eingereicht und sich noch im Bewertungsverfahren befindet. Sie ist entstanden in Zusammenarbeit mit Prof. Dr. Uta Störl, FernUniversität in Hagen, und Prof. Dr.-Ing. Meike Klettke, Universität Regensburg.

Literatur

- [1] K. Biernacka und S. Schulz. *Forschungsdatenmanagement in der Informatik*. Humboldt-Universität zu Berlin, 2022.
- [2] S. Diemt. A Requirements Review Based on a User Survey for a University-Wide User-Centered Research Data Management System. (eingereicht zur Begutachtung), 2024.
- [3] M. Prensky. Digital Natives, Digital Immigrants Part 1. *On the Horizon*, 9(5):1–6, 2001.
- [4] M. D. Wilkinson et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3:160018, 2016.