

## Fachbeitrag

Heike Neuroth, Claudia Engelhardt, Jochen Klar, Jens Ludwig und Harry Enke

# Aktives Forschungsdatenmanagement

<http://doi.org/10.1515/abitech-2018-0008>

**Zusammenfassung:** Forschungsdatenmanagement und damit einhergehend Forschungsdatenmanagementpläne nehmen national und international an Bedeutung zu. Nicht nur, dass verschiedene Förderorganisationen wie die National Science Foundation (USA), der Schweizerische Nationalfonds (SNF), die Deutsche Forschungsgemeinschaft (DFG) sowie die Europäische Kommission mit Horizon 2020 (H2020) bereits bei Projektanträgen Auskunft über den Umgang mit den nachgenutzten oder erstellten Forschungsdaten verlangen, es beschäftigen sich auch mehr und mehr Initiativen wie zum Beispiel international die Research Data Alliance oder in Deutschland die DINI/nestor-Arbeitsgruppe Forschungsdaten mit dem Thema. International setzt sich dabei mehr und mehr die Erkenntnis durch, dass es im Umgang mit Forschungsdaten nicht mit einem einmaligen Erstellen eines Forschungsdatenmanagementplans getan ist, sondern dass sich die Pläne aktiv dem Verlauf des Forschungsprozesses anpassen und für verschiedene Bedarfe zur Verfügung gestellt werden müssen. So kann es sinnvoll und notwendig sein, weitere Beteiligte wie zum Beispiel IT-Support oder übergeordnete Datenmanager beim Erstellen und Aktualisieren eines Forschungsdatenmanagementplans zu berücksichtigen. Daher muss ein Werkzeug zur Unterstützung von diesen Plänen über das bloße Ausfüllen von Vorlagen der Förderorganisation hinaus weitere Aufgaben erfüllen und so den gesamten Prozess des Forschungsdatenmanagements unterstützen. Der Research Data Management Organiser (RDMO) ist ein solches Werkzeug, das im Rahmen eines DFG-Projektes entwickelt und mit Hilfe unterschiedlicher Gruppen von Nutzenden getestet wurde. Das RDMO-Tool ist multilingual, flexibel an Community- und Organisationsanforderungen anpassbar und unterstützt verschiedene Aufgaben wie zum Beispiel unterschiedliche Export-Funktionen oder die Erledigung zeitlich gebundener Tasks. In einem DFG-Nachfolgeprojekt, das Ende 2017 gestartet ist, werden in RDMO wichtige Erweiterungen sowohl technischer Art als auch bezogen auf verschiedene Aspekte der Nachhaltigkeit vorgenommen.

**Schlüsselwörter:** aktives Forschungsdatenmanagement, Forschungsdatenmanagementplan, Dokumentation von Forschungsdaten

## Research Data Management Organiser (RDMO)

**Abstract:** Research data management gets more and more international attention, and one of its major components are research data management plans. Various funding agencies, e.g. the National Science Foundation in the US, the Schweizer Nationalfonds (SNF), the Deutsche Forschungsgemeinschaft (DFG), and very prominently the European Commission with Horizon 2020 (H2020) ask for concise descriptions of data management as part of a funding proposal. Also, initiatives like the Research Data Alliance (RDA) or the DINI/nestor workgroup Research Data are focusing attention on this topic. Management of research data is no longer viewed as done by drafting merely a static research data management plan, but is now regarded as a process of active support and care during the whole research process. Different participants of the research process need different and up-to-date status information. A tool to support data management during the whole research process has more to deliver than a plain document as required by the respective funding agencies. The Research Data Management Organiser (RDMO) is a companion tool for organising this process as a collaborative effort of various stakeholders or participants. RDMO features multilingual capabilities, adaptability to requirements of communities or organisations, a templating system for export of collected data, snapshotting and timeline capabilities. A second project phase, also funded by DFG, has started end of 2017 to extend RDMO's technical capabilities and ease of use, and to address the aspects of sustainability and community support.

**Keywords:** active research data management, research data management plan, documentation of research data

## 1 Einleitung

### 1.1 Forschungsdaten(management)

Die Notwendigkeit des Forschungsdatenmanagements (FDM) in wissenschaftlichen und zum Teil auch in wirt-

schaftlichen<sup>1</sup> Vorhaben sowie der Wert von Forschungsdaten sind in den letzten Jahren wesentlich deutlicher in das Bewusstsein sowohl der Forschungsorganisationen als auch der Forschungsförderer gerückt. Eine der wichtigsten Komponenten des Forschungsdatenmanagements sind Datenmanagementpläne (DMP), deren Nutzung und Potential jedoch weiterhin nicht ausgeschöpft sind. Auf Seiten der Forschungsförderer hat dies vor allem dazu geführt, die aus dem angloamerikanischen Raum stammenden DMP in den Förderbedingungen zu verankern. Hier ist insbesondere die Europäische Kommission mit dem Förderprogramm Horizon 2020 zu nennen. Der zunächst auf sieben, später auf neun Forschungsgebiete angelegte Open Research Data Pilot<sup>2</sup> wurde ab 2017 auf das gesamte Förderportfolio ausgedehnt.<sup>3</sup> Das bedeutet konkret, dass jedes Horizon 2020-Projekt, das nicht von vornherein eine begründete Opt-Out-Möglichkeit<sup>4</sup> wählt, einen DMP erstellen muss. In den Horizon 2020-Richtlinien zum Datenmanagement wird nunmehr explizit betont, dass Datenmanagementpläne im Projektverlauf zu aktualisieren sind, wenn sich signifikante Änderungen in Bezug auf die Daten ergeben. Dies solle unabhängig von der regulären Kommunikation mit dem Förderer geschehen. Die Aktualisierung bezieht sich nicht nur auf neu hinzugekommene Datensätze während des Projektverlaufs, sondern beispielsweise auch auf Änderungen in der Projektzusammensetzung oder Ergänzungen von relevanten Leitlinien, Policies usw. Es wurde erkannt, dass statische DMP, also einmal zu Beginn eines Projektes angefertigte und danach nicht mehr aktualisierte Pläne, wenig hilfreich sind für die Dokumentation und v. a. Interpretation der Forschungsdaten. Vielmehr wird der Prozess des Forschungsdatenmanagements jetzt umfassender gesehen und unterliegt

deutlich dynamischen Verläufen. Dies führt letztendlich hin zu einem aktiven DMP, der über die ganze Laufzeit eines Projektes gepflegt und aktualisiert wird. Solche aktiven DMP reflektieren immer den aktuellen Stand aller für das Datenmanagement wichtigen Aspekte und sind eine relevante Informationsquelle nicht nur für die am Projekt beteiligten Akteure, sondern auch für die Nachnutzung der Forschungsdaten.

In Deutschland veröffentlichte die DFG 2015 ihre Leitlinien zum Umgang mit Forschungsdaten.<sup>5</sup> In diesen werden Datenmanagementpläne zwar (noch) nicht explizit erwähnt, jedoch werden Antragstellende dazu aufgefordert, in ihrem Antrag Aussagen zu Erzeugung, Auswertung usw. von Forschungsdaten zu tätigen. Dies umfasst zum Beispiel auch Informationen zu den verwendeten Datentypen, zur Nutzung fachspezifischer Standards und zur Wahl angemessener Repositorien, zu den Rechten betroffener Dritter sowie zu den Planungen bezüglich der Datenveröffentlichung – und somit Elemente, die typischerweise in einem Datenmanagementplan enthalten sind. Immer mehr wissenschaftliche Einrichtungen und Infrastrukturanbieter greifen das Thema auf, etwa indem sie Forschungsdaten-Policies<sup>6</sup> verabschieden und Unterstützungsangebote zum Forschungsdatenmanagement für Wissenschaftlerinnen und Wissenschaftler<sup>7</sup> aufbauen. Dabei möchten viele Hochschulen und Wissenschaftseinrichtungen ihren Forschenden nicht nur Beratung anbieten, sondern zum Teil auch Software-Werkzeuge zur Verfügung stellen, die nicht nur die Erstellung von Datenmanagementplänen erleichtern, sondern auch die Planung und Organisation des Datenmanagements projektbegleitend unterstützen. Die beiden international verbreiteten Software-Werkzeuge DMPTool<sup>8</sup> und DMPonline<sup>9</sup> berichten von gemeinsam mehr als 21 000 Nutzenden und 22 500 DMP bis Ende 2015.<sup>10</sup> Beide Tools unterstützen aber bisher stärker Aspekte wie Controlling und projektbegleitende Aktualisierung von Datenmanagementplänen als die konkrete Aufgabenorganisation und verantwortliche Rollenverteilung im Pro-

1 So veröffentlichte z. B. im August 2017 das Bundesministerium für Verkehr und digitale Infrastruktur (BMVI) die Ausschreibung „Angewandte Forschung und Experimentelle Entwicklung“, in der explizit als Kriterium genannt wurde: „Ein Einverständnis zu nachfolgender Datennutzungsregelung ist zwingende Voraussetzung für die Skizzenreichung.“ Dort heißt es u. a.: „Die Bereitstellung erster Daten hat spätestens 1 Jahr nach Laufzeitbeginn stattzufinden und ist für die Dauer von mindestens 3 Jahren nach Laufzeitende zu gewährleisten.“ Vgl. [http://www.bmvi.de/SharedDocs/DE/Anlage/DG/mfund-dritter-foerderaufruf-foerderlinie-zwei.pdf?\\_\\_blob=publicationFile](http://www.bmvi.de/SharedDocs/DE/Anlage/DG/mfund-dritter-foerderaufruf-foerderlinie-zwei.pdf?__blob=publicationFile) (02.02.2018).

2 Vgl. <https://www.openaire.eu/opendatapilot> (02.02.2018).

3 Vgl. European Commission. *Guidelines on FAIR Data Management in Horizon 2020*, Version 3.0, 26 Juli 2016, 3. [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf) (02.02.2018).

4 Vgl. z. B. Österreichische Forschungsförderungsgesellschaft (FFG): *Open Data in Horizon 2020*. [https://www.ffg.at/europa/recht-finanzen/h2020-open\\_data](https://www.ffg.at/europa/recht-finanzen/h2020-open_data) (02.02.2018).

5 Deutsche Forschungsgemeinschaft (DFG). *Leitlinien zum Umgang mit Forschungsdaten*. 30.09.2015. [http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien\\_forschungsdaten.pdf](http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf) (02.02.2018).

6 Vgl. z. B. [http://www.forschungsdaten.org/index.php/Data\\_Policies](http://www.forschungsdaten.org/index.php/Data_Policies) (02.02.2018).

7 Vgl. z. B. <https://www.zotero.org/groups/439756/forschungsdaten/items/collectionKey/5MK88CAE> (02.02.2018).

8 DMPTool. <https://dmptool.org> (02.02.2018).

9 DMPonline. <https://dmponline.dcc.ac.uk> (02.02.2018).

10 Vgl. Simms, Stephanie, Marisa Strong, Sarah Jones et al. „The Future of Data Management Planning: Tools, Policies, and Players“. *International Journal of Digital Curation* 11,1 (2016): 208–217, hier: 211. DOI:10.2218/ijdc.v11i1.413.

jekt. In einer gemeinsamen DMPRoadmap<sup>11</sup> formulieren sie als Hauptziele für zukünftige Entwicklungen Internationalisierung, Nachnutzung von Metadaten, Offenheit und Evaluierung (Assessment). In dieser Diskussion spielt die geplante Unterstützung von aktualisierbaren DMP (Active DMPs) eine wichtige Rolle.<sup>12</sup> Eine neuere Überlegung ist es, DMP auch maschinell ausführbar zu machen (Actionable DMPs) und für die Automatisierung und Kontrolle des Datenmanagements zu nutzen, wie der Zusammenschluss im Rahmen der DMPRoadmap dokumentiert. Es wird jedoch zugleich angemerkt, dass bisher nur wenige praktikable Resultate vorliegen.<sup>13</sup> In DMP-Roadmap wird, im Gegensatz zu RDMO, weiterhin an einer zentralen (nationalen) Instanz für die Ablage von DMP festgehalten.

## 1.2 Hintergrund des DFG-Projekts

Die DFG förderte ab November 2015 für 18 Monate das Vorhaben „Entwicklung und Implementierung eines Werkzeugs für die Planung, Umsetzung und Kontrolle des Forschungsdatenmanagements (FDMP-Werkzeug)“ und fördert seit November 2017 für 30 Monate das Nachfolgeprojekt „Research Data Management Organiser (RDMO)“. Während die erste Phase von den beiden Projektpartnern Leibniz-Institut für Astrophysik Potsdam (AIP) und Fachbereich Informationswissenschaften an der Fachhochschule Potsdam (FHP) verantwortet wurde, ist in der zweiten Phase die Bibliothek des Karlsruher Instituts für Technologie (KIT) zum Projektkonsortium hinzugestoßen. Mit Ende der ersten Förderphase steht nun ein Werkzeug zur Verfügung, das die strukturierte Planung, Umsetzung und Verwaltung des Forschungsdatenmanagements unterstützt und zusätzlich die textuelle Ausgabe von Datenmanagementplänen ermöglicht. Konzeptuell ist die Begleitung des gesamten Forschungsprozesses mit Hilfe des Werkzeugs möglich. Im Verlauf der ersten Projektphase wurde das Werkzeug in Research Data Management Organiser (RDMO) umbenannt, um den aktiven Charakter von Forschungsdatenmanagement zu betonen und darauf aufmerksam zu machen, dass verschiedene Stakeholder mit zum Teil ganz unterschiedlichen Perspektiven das Werkzeug nutzen können. RDMO ermöglicht es den Forschern

den, aber auch anderen am Forschungsprozess beteiligten Akteuren wie IT-Verantwortlichen, Bibliothek oder Rechenzentrum, alle zum Forschungsdatenmanagement nötigen Informationen an einem Punkt einzugeben und in einem kollaborativen Prozess über die ganze Projektlaufzeit aktuell zu halten („Active DMP“). Die geschieht in Form eines strukturierten Online-Interviews auf Basis eines konfigurierbaren Fragenkatalogs. Auf die Informationen kann dann auf verschiedene Weise zugegriffen werden. So können für die Antragstellung oder Berichte textuelle DMP nach hinterlegten Vorlagen (beispielsweise H2020 DMP) erstellt und in gängigen Formaten exportiert werden. Über eine API können die Daten in weiteren Software-Komponenten weiterverarbeitet werden.

Die Software kann, mit überschaubarem Aufwand, von einzelnen Forschungseinrichtungen und Hochschulen, aber auch Kollaborationen oder Projektverbänden, eigenständig installiert und betrieben werden. Hierbei können eigene Fragenkataloge und Vorlagen hinterlegt oder der innerhalb des RDMO-Projekts konzipierte generische Fragenkatalog angepasst werden. Hierdurch wird eine optimale Integration einer eigenen RDMO-Instanz in den fachlichen oder lokalen Kontext des Betreibers ermöglicht. Die eingegebenen Daten verbleiben beim Betreiber und müssen nicht an Dritte weitergegeben werden.

## 2 Research Data Management Organiser (RDMO)

Die Aufgaben und Arbeiten im Projekt waren in drei Arbeitspakete (AP) gegliedert:

- AP 1 Konzeption: Dieses Arbeitspaket widmete sich der Konzeption der Inhalte, der Struktur und des Bedienungsablaufs des im Tool verwendeten Fragenkatalogs und untersuchte den disziplinspezifischen Anpassungsbedarf am Beispiel der Astrophysik und der Sozialwissenschaften.
- AP 2 Software: Aufgabe dieses Arbeitspaketes war die technische Entwicklung des Tools, was auch die Konzeption der Softwarearchitektur und des Datenmodells beinhaltete.
- AP 3 Community und Nutzertests: Schwerpunkt dieses Arbeitspaketes war der Austausch mit den Zielgruppen über verschiedene Kanäle bezüglich der Anforderungserhebung und die Evaluierung des RDMO-Tools.

Im Folgenden werden die Aktivitäten und Ergebnisse in diesen Bereichen beschrieben, wobei den Arbeiten in AP 3 aufgrund ihrer engen Verquickung mit denen in den ande-

<sup>11</sup> <https://github.com/dmproadmap> (02.02.2018).

<sup>12</sup> Vgl. Simms, Stephanie, Sarah Jones, Kevin Ashley et al. „Roadmap: A Research Data Management Advisory Platform.“ *Research Ideas and Outcomes* 2 (2016): e8649. DOI:10.3897/rio.2.e8649. „Active Data Management Plans“ werden auch in einer gleichnamigen RDA Interest Group behandelt, siehe: <https://www.rd-alliance.org/groups/active-data-management-plans.html> (02.02.2018).

<sup>13</sup> Vgl. Simms/Jones/Ashley et al. 2016.

ren beiden APs kein eigenes Unterkapitel gewidmet, sondern über sie im Zusammenhang mit diesen berichtet wird.

## 2.1 Konzeption

Die Konzeption des vom RDMO-Projekt entwickelten Fragenkatalogs baut auf eigenen Vorarbeiten aus früheren Projekten auf. Hier ist insbesondere der „Leitfaden zum Forschungsdatenmanagement“<sup>14</sup> zu nennen, der im Projekt WissGrid (2009–2012) entwickelt wurde und auf Vorarbeiten insbesondere des Digital Curation Center (DCC)<sup>15</sup> aufbaut, die basierend auf eigener Expertise ergänzt, modifiziert und strukturiert wurden. Ein weiterer wichtiger Baustein war eine Weiterentwicklung der WissGrid-Checkliste, die am Max-Planck-Institut zur Erforschung multireligiöser und multiethnischer Gesellschaften (MPI MMG) im Rahmen eines Projektes mit der Niedersächsischen Staats- und Universitätsbibliothek Göttingen erarbeitete wurde und in der bereits Folgeaufgaben zu Fragen definiert wurden. Neben diesen beiden Hauptquellen wurden auch Datenmanagementpläne und Checklisten anderer Tools und Institutionen sowie Vorgaben oder Richtlinien relevanter Forschungsförderer evaluiert, aus denen ebenfalls Anregungen aufgenommen wurden. Nennenswert sind hier vor allem die Checkliste der Universität Bielefeld<sup>16</sup> und die Guidelines on FAIR Data Management in Horizon 2020.<sup>17</sup> Auf Grundlage dieser Vorarbeiten wurde ein Entwurf der Struktur und Abfolge der Fragen erarbeitet und in einem ersten Prototyp umgesetzt, der im Rahmen mehrerer Experteninterviews mit verschiedenen Zielgruppen<sup>18</sup> vorgestellt wurde. Auf Basis des Feedbacks zum Fragenkatalog, den Funktionalitäten und der Usability wurden diese iterativ weiterentwickelt. Des Weiteren wurden auch aktuelle Entwicklungen und Bedarfe<sup>19</sup> einbezogen. Auf Expertenworkshops mit Vertreterinnen und Vertretern von Infrastrukturu-

reinrichtungen und -initiativen sowohl mit institutioneller als auch mit fachwissenschaftlicher Ausrichtung wurden ebenfalls Anregungen und Bedarfe gesammelt, die in die Weiterentwicklung von RDMO geflossen sind.

Zur Definition der Anforderungen verschiedener Stakeholder in Bezug auf inhaltliche Aspekte wie auch die Usability wurde der aus der agilen Softwareentwicklung stammende Ansatz der User Stories<sup>20</sup> genutzt. Dabei werden die Anforderungen in strukturierter Form beschrieben, die die Perspektive unterschiedlicher Akteursgruppen („In meiner Rolle als A ...“), die gewünschte Interaktion zwischen Akteur und System („... möchte ich B tun ...“) und den beabsichtigen Nutzen bzw. das Ziel der Interaktion („..., um den Nutzen C zu erreichen.“) einbezieht.

Es wurden insgesamt etwa 70 User Stories für die folgenden Akteursgruppen (bzw. Rollen)<sup>21</sup> definiert:

- Autorin oder Autor,
- Infrastrukturdienstleister,
- Vorgesetzte oder Vorgesetzter bzw. Institutsleitung,
- Datenmanagerin oder Datenmanager,
- Gast,
- Managerin oder Manager,
- IT-Administratorin oder IT-Administrator bzw. IT-Support und
- Entwicklerin oder Entwickler.

Die folgenden Kategorien bilden die Bereiche ab, die für das Erreichen des Ziels einer Interaktion verantwortlich sind:

- Versionierung,
- Kollaboration,
- Usability, Eingabehilfen und Vorlagen,
- Anpassbarkeit der Fragen und Antwortoptionen,
- Fragenlogik,
- Aufgabenverwaltung,
- Rollendefinition und
- Schnittstellen und Exportmöglichkeiten.

Im Folgenden finden sich einige Beispiele für die User Stories, jeweils charakterisiert durch die betreffende Akteursgruppe, die gewünschte Interaktion, den erhofften Nutzen und die relevante Kategorie:

- Als Autorin oder Autor möchte ich auf alte Versionen meines DMP zugreifen, um fehlerhafte Bearbeitun-

<sup>14</sup> Ludwig, Jens, Harry Enke (Hrsg.). *Leitfaden zum Forschungsdaten-Management*. Glückstadt: Verlag Werner Hülsbusch, 2013.

<sup>15</sup> Donnelly, Martin, Sarah Jones. *Checklist for a Data Management Plan*. Version 3.0, 17. März 2011. O.O.: DCC, 2011. [http://www.dcc.ac.uk/sites/default/files/documents/data-forum/documents/docs/DCC\\_Checklist\\_DMP\\_v3.pdf](http://www.dcc.ac.uk/sites/default/files/documents/data-forum/documents/docs/DCC_Checklist_DMP_v3.pdf) (02.02.2018).

<sup>16</sup> <https://data.uni-bielefeld.de/de/data-management-plan> (02.02.2018).

<sup>17</sup> European Commission. *Guidelines on FAIR Data Management in Horizon 2020*. Version 3.0, 26 Juli 2016. [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf) (02.02.2018).

<sup>18</sup> Z. B. den Arbeitsgruppen TwinLife (<http://www.twin-life.de/>, 02.02.2018) und Sonnenphysik (AIP Potsdam).

<sup>19</sup> Insbesondere die im Juli 2016 veröffentlichten „Guidelines on FAIR Data Management in Horizon 2020“.

<sup>20</sup> Vgl. z. B. Ambler, Scott W. „User Stories. An Agile Introduction.“ In *Agile Modeling*. Ambysoft Inc. (Hg.). 2014. <http://www.agilemodeling.com/artifacts/userStory.htm> (02.02.2018).

<sup>21</sup> Im weiteren Verlauf des Projektes muss geprüft werden, inwieweit andere Rollen wie zum Beispiel Forschungsreferentin oder -referent sowie Gutachterin oder Gutachter bzw. Förderer (s. u.) ebenfalls zu berücksichtigen sind.

- gen nachvollziehen zu können (Kategorie: Versionierung).
- Als Autorin oder Autor möchte ich andere Personen zu meinem DMP als Leserin bzw. Leser oder Autorin bzw. Autor einladen können, damit sie mitwirken können (Kategorie: Kollaboration).
- Als Autorin oder Autor möchte ich Vorlagen meiner Institution und Förderer sowie Empfehlungen für DMP nutzen, um den Schreibprozess zu beschleunigen (Kategorie: Usability, Eingabehilfen, Vorlagen).
- Als Infrastrukturdienstleister möchte ich Auswahlfelder mit praktikablen, korrekten und standardisierten Alternativen für möglichst viele Fragen definieren können, damit ich die Nutzerantworten standardisieren und einfacher auswerten kann (Kategorie: Anpassbarkeit der Fragen und Antwortoptionen).
- Als Datenmanagerin oder -manager<sup>22</sup> möchte ich Antworten aus ganz verschiedenen RDMO-Installationen abfragen können, um Statistiken über verschiedene Einrichtungen hinweg erstellen zu können (Kategorie: Schnittstellen und Exportmöglichkeiten).
- Als Infrastrukturdienstleister möchte ich die Antworten eines DMP einsehen können, um entsprechende Planungen vornehmen zu können, etwa zur rechtzeitigen Bereitstellung von Speicherplatz (Kategorie: Aufgabenverwaltung).
- Als Managerin oder Manager möchte ich im Auftrag des Betreibers eine Instanz von RDMO inhaltlich und organisatorisch betreuen, um die Nutzenden und den Betreiber bei der Planung des DMP zu unterstützen (Kategorie: Rollendefinition).
- Als Vorgesetzte oder Vorgesetzter bzw. Institutionsleitung möchte ich eine graphische Übersicht über die Daten in RDMO abrufen können, um vergleichen und kontrollieren zu können (Kategorie: Usability, Eingabehilfen, Vorlagen).
- Als IT-Administratorin oder -Administrator bzw. IT-Support möchte ich im Auftrag des Betreibers eine Instanz von RDMO technisch betreuen, um die Verlässlichkeit und den Betrieb der Services sicherzustellen (Kategorie: Rollenverteilung).
- Als Entwicklerin oder Entwickler möchte ich die technische Plattform RDMO weiterentwickeln oder anpassen, um das Werkzeug nachnutzen zu können und

nicht etwas komplett Neues entwickeln zu müssen (Kategorie: Rollendefinition).

- Als Infrastrukturdienstleister möchte ich Fragen und Antworten verschiedener Granularität anbieten, damit je nach Bedarf, Fähigkeit und Bereitschaft im praktikablen Detailgrad geplant werden kann (Kategorie: Fragenlogik).

Die meisten Rollen in den User Stories sind bisher mit der Akteursgruppe „Autorin oder Autor“ verknüpft (über die Hälfte), während die Kategorien mehr oder weniger gleichmäßig verteilt sind. Es wird interessant sein zu sehen, welche Akteursgruppen und Kategorien bei einem breiten Einsatz des RDMO-Werkzeugs noch zum Einsatz kommen werden. Eine finale Definition der Akteursgruppen wird in einem nächsten Schritt, abgestimmt mit den bisherigen Interessenten, entwickelt. Es dürfte nicht ganz einfach sein, die notwendigen unterschiedlichen Rollen auf die verschiedenen organisatorischen Strukturen der Forschungseinrichtungen und Hochschulen adäquat abzubilden.

Ausgehend von den oben genannten Vorarbeiten wurde zu Beginn des Projektes eine erste Entwurfsversion der Fragenstruktur erstellt. Der Fragenverlauf einschließlich der Abhängigkeiten wurde zunächst mit Hilfe des yED Graph Editors<sup>23</sup> visualisiert. Diese erste Version wurde in den Prototyp des Werkzeugs implementiert und projektintern getestet. Hierbei wurden zu einer Vielzahl von Fragen auch bereits Hilfetexte und externe Links eingefügt sowie bereits die Übersetzung der Fragen, Antworten und Hinweistexte ins Englische vorgenommen. Ab diesem Zeitpunkt konnten weitere Änderungen an den Fragen dann auch direkt im Werkzeug umgesetzt werden.

Die Veröffentlichung der neuen Richtlinien der Europäischen Kommission zum FAIR Data Management in Horizon 2020, die wesentlich detaillierter und umfangreicher sind als die zuvor geltenden, führte ebenfalls noch einmal zu einer Überarbeitung des Fragenkatalogs. So ist eine Reihe von Fragen neu aufgenommen worden, um mit den H2020-Anforderungen vollständig kompatibel zu sein.

Die zum Projektende der ersten Förderphase vorliegende Version des Fragenkatalogs umfasst sieben Themengebiete, die sich wiederum in Unterthemen und schließlich in einzelne Fragen oder Fragensets aufgliedern:

- **Allgemeine Angaben zum Projekt:** In diesem Abschnitt werden wichtige Rahmendaten eines Projekts erhoben. Dies umfasst z. B. die Forschungsfrage, die Projektlaufzeit, die Projektpartner einschließlich der bei diesen jeweils für das Datenmanagement verant-

<sup>22</sup> Dies könnte auch für eine potentielle Rolle Gutachterin bzw. Gutachter oder Förderer interessant sein: Als Gutachter/Förderer möchte ich Zugang zu den Daten in RDMO haben, um die angemessene Planung und Umsetzung des Datenmanagements prüfen zu können; dies betrifft die Kategorie: Rollendefinition.

<sup>23</sup> <https://www.yworks.com/products/yed> (02.02.2018).

wortlichen Personen sowie Richtlinien im Umgang mit Forschungsdaten, etwa der eigenen oder von Partnerinstitutionen, dem Forschungsförderer oder von Fachgesellschaften, die berücksichtigt werden sollten.

- **Inhaltliche Einordnung der Daten:** In diesem Abschnitt wird beschrieben, welche Art bzw. welche Arten von Daten genutzt werden, ob sie nachgenutzt oder selbst erhoben bzw. selbst erstellt werden sowie ob, und, wenn ja, mit welchem Aufwand sie reproduzierbar sind. Zudem wird gefragt, für wen die Daten für die Nachnutzung eventuell interessant sein könnten.
- **Technische Einordnung der Daten:** In diesem Themenblock werden eine Reihe von Informationen erfasst, die zum einen für die Speicherung und Nutzung der Daten während der Projektlaufzeit, zum anderen aber auch für die Langzeitarchivierung und die Nachnutzung von großer Bedeutung sind. Hierzu gehören die Größe bzw. Menge der Daten, die Formate und die Programme und Technologien, die zur Erzeugung der Daten genutzt werden. Auch werden mit den Zeitpunkten der Datenerhebung, -aufbereitung, -analyse und -veröffentlichung wichtige Zeitpunkte im Datenlebenszyklus erhoben.
- **Datennutzung während des Projekts:** Die Fragen in diesem Abschnitt haben die Datennutzung während der Laufzeit eines Vorhabens im Fokus. Ziel ist es, anhand der gewonnenen Informationen einschätzen zu können, welche Infrastrukturressourcen und ggf. auch, welche Unterstützung durch IT- oder (zentrales) Datenmanagement benötigt werden. Konkret werden Angaben zur Datenspeicherung und -sicherheit, zum Teilen und der Veröffentlichung, zum kollaborativen Arbeiten an den Daten, zu unterstützungsbedürftigen Nutzungsszenarien, zur Qualitätssicherung und zu den Kosten verschiedener Aspekte oder Schritte der Datennutzung erhoben. Bezüglich der Kosten wird unterschieden zwischen Personalkosten, die in Personenmonaten anzugeben sind, und Sachkosten, die in Euro erfragt werden. Dabei wird es sich, insbesondere in der Planungsphase oder zu Beginn eines Projektes, i. d. R. um Schätzungen handeln. In den folgenden Themenblöcken werden ebenfalls die Aufwände für jeweils relevante Maßnahmen in dieser Form erhoben. Werden die einzelnen Posten zusammengezählt, kann am Ende eine Schätzung des für das Datenmanagement im Projekt erforderlichen Gesamtbedarfs an Personal- und Sachmitteln ermittelt werden.
- **Metadaten und Referenzierung:** Mit Metadaten werden Daten beschrieben, damit sie verstanden, sinnvoll verwaltet und später gezielt wiedergefunden werden können. Dies ist sowohl für die interne Organisation

und Verwaltung der Daten während des Projektes als auch für die Langzeitarchivierung, Veröffentlichung und Nachnutzung wichtig. Auch das Forschungsdezernat oder die Leitung einer Institution, die eine RDMO-Instanz betreibt und einen Überblick über alle an der Einrichtung genutzten bzw. erzeugten Daten erlangen will, ist auf diese Metadaten entsprechend angewiesen. In diesem Block wird auch erfragt, welche Metadaten notwendig sind, um die Daten und ihre Genese zu verstehen und welche davon manuell, automatisch oder semi-automatisch erhoben werden. Zudem wird nach der Nutzung von persistenten Identifikatoren gefragt, die eine dauerhafte und verlässliche Referenzierung digitaler Objekte ermöglichen.

- **Rechtliche und ethische Fragen:** In Bezug auf rechtliche und ethische Fragen im Zusammenhang mit Forschungsdaten sind insbesondere zwei Gebiete relevant: sensible Daten und urheberrechtlich oder durch verwandte Schutzrechte geschützte Daten. Bei sensiblen Daten spielen personenbezogene Daten die größte Rolle, doch können Daten auch aus anderen Gründen sensibel sein (z. B., weil sie Standortinformationen zu geschützten Pflanzen- oder Tierarten oder Betriebsgeheimnisse enthalten). Geben Nutzen an, mit personenbezogenen Daten zu arbeiten, folgen Fragen zu Datenschutz und Anonymisierung. Außerdem wird erhoben, ob das Vorhaben von einer Ethikkommission begutachtet worden ist und ob behördliche Genehmigungen benötigt werden.
- **Langfristige Archivierungsstrategie:** Der letzte Themenblock widmet sich der Langzeitarchivierung der Forschungsdaten nach dem Projektende. Hier werden Fragen zur Auswahl der zu archivierenden Daten gestellt. Dazu gehören Fragen wie aus welchen Gründen, nach welchen Kriterien und durch wen die Auswahl erfolgt. Es wird außerdem erhoben, wann, wie lange und wo die Daten gespeichert werden sollen sowie, ob das ausgewählte (fachspezifische) Repositorium oder Datenzentrum zertifiziert ist und mit ihm bereits Kontakt aufgenommen wurde, um die Voraussetzungen und Rahmenbedingungen der Datenübernahme zu klären. Auch wird gefragt, ob die Daten erst nach einer Sperrfrist zugänglich gemacht werden sollen.

Auf Seiten der Antwortmöglichkeiten wurde soweit möglich auf die Nutzung von Freitextfeldern verzichtet, da so die Antworten, insbesondere über verschiedene Projekte hinweg, besser vergleichbar und weiter (automatisch) verarbeitbar sind. Auch stellt dies eine Voraussetzung dar, um bestimmte Aufgaben automatisiert zu implementieren, die dann relevante Stakeholder zu einem definierten

Zeitpunkt über weitere To-dos alarmieren. Für einen Großteil der relevanten Fragen konnte dies umgesetzt werden. Dabei wurden, wo es sinnvoll schien, kontrollierte Vokabulare eingebunden.

Der Fokus bei der Konzeption des detaillierten Fragenkatalogs lag darauf, alle Aspekte zu berücksichtigen, die für das Datenmanagement notwendig sind. Der Fragenkatalog wird im RDMO-eigenen XML-basierten Austauschformat auf GitHub zur Verfügung gestellt<sup>24</sup> und kann von Institutionen, Projekten usw. in eigenen RDMO-Instanzen nachgenutzt werden. Er ist als Baukasten oder Anregung für die Erstellung eines auf die Bedürfnisse des jeweiligen Kontextes angepassten, eigenen Fragenkataloges gedacht.

Teil der konzeptionellen Arbeit bezüglich des Fragenkatalogs war die Untersuchung des fachspezifischen Anpassungsbedarfs am Beispiel der Sozialwissenschaften und der Astrophysik. Hierfür wurde zunächst die Fachliteratur<sup>25</sup> ausgewertet. Die so gewonnenen Informationen wurden in den Interviews mit Vertreterinnen und Vertretern der Beispieldisziplinen überprüft sowie anschließend auch noch einmal im Rahmen von Expertenworkshops zur Diskussion gestellt. Dabei hat sich gezeigt, dass der Anteil an Fragen, für die ein disziplinspezifischer Detailgrad sinnvoll wäre,<sup>26</sup> nur sehr gering ist. Eine weitergehende Anpassung ist erst auf einer sehr feingranularen Ebene sinnvoll. Dies führt dann aber dazu, dass der resultierende Fragenkatalog so eng auf einen sehr eingeschränkten Nutzendenkreis zugeschnitten wäre, dass sich ein solches Vorgehen wahrscheinlich nur in wenigen speziellen Fällen empfiehlt.

Bei den Antwortoptionen sowie für die Hilfe- bzw. Hinweistexte ergibt sich dagegen ein anderes Bild: Hier

wird eine disziplin- ggf. auch methodenspezifische Anpassung – etwa indem im Antwortbereich Dropdown-Listen mit einer Vorauswahl an jeweils relevanten disziplinspezifischen Optionen angeboten werden – als sehr nützlich angesehen. Gleiches gilt für Hinweistexte, die mit passenden Beispielen oder mit Verweisen auf für die jeweilige Disziplin relevante Ressourcen angereichert werden können. Die Teilnehmerinnen und Teilnehmer des Expertenworkshops waren sich einig, dass eine solche Anpassung ausgesprochen vorteilhaft wäre. Dies stelle zum einen für die jeweilige Fachdisziplin eine große Hilfe beim Beantworten der Fragen und der darauf aufbauenden Verwaltung und Analyse relevanter Informationen dar, zum anderen erleichtert es durch die Standardisierung eines möglichst großen Teils der Antworten die Aggregation und das Vergleichen von Informationen über mehrere Projekte hinweg.

Jedoch kann eine solche Anpassung nicht für eine Vielzahl von Disziplinen zentral von einem Projekt wie RDMO geleistet werden, sondern muss vielmehr durch die Fachwissenschaften (z. B. organisiert in Fachgesellschaften) selbst erfolgen. Der Beitrag von RDMO ist ein generischer Fragenkatalog als Grundlage und Rahmen hierfür sowie die Bereitstellung eines technischen Werkzeuges, das es erlaubt, Fragenkatalog und Hilfetexte flexibel und einfach zu editieren und dynamisch an die Erfordernisse verschiedener Kontexte anzupassen. Zudem kann das RDMO-Projekt die nötigen Prozesse moderieren und verschiedene interessierte Fach-Communities zusammenbringen. Interessant in diesem Zusammenhang ist, dass sich eine Interest Group (IG) bei der internationalen Initiative Research Data Alliance seit kurzer Zeit genau dieser Herausforderung stellt. Die IG „Domain Repositories“<sup>27</sup> beschäftigt sich mit der Entwicklung von domainspezifischen Datenmanagementplänen, die zu einer internationalen Akzeptanz, Nachnutzung und Verbreitung in den jeweiligen Fachdisziplinen führen soll.

## 2.2 Software

Das RDMO-Werkzeug ist als interaktive Webanwendung realisiert und steht als Open-Source unter der Apache-2.0-Lizenz zur Verfügung.<sup>28</sup> Die Software ermöglicht es sowohl einzelnen Forschenden als auch Kollaborationen, alle für das nachhaltige Datenmanagement nötigen Informationen zentral einzugeben und zu verwalten. Auf der ober-

<sup>24</sup> <https://github.com/rdmorganiser/rdm-catalog> (02.02.2018).

<sup>25</sup> Vor allem Enke, Harry, Joachim Wambsgang. „Astronomie und Astrophysik.“ In *Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme*. Hrsg. von Heike Neuroth et al. Boizenburg: Verlag Werner Hülsbusch, 2012, 275–293. Auch: [http://nestor.sub.uni-goettingen.de/bestandsaufnahme/kapitel/nestor\\_bestandsaufnahme\\_014.pdf](http://nestor.sub.uni-goettingen.de/bestandsaufnahme/kapitel/nestor_bestandsaufnahme_014.pdf); Häder, Michael. *Der Datenschutz in den Sozialwissenschaften*. 2009. (RatSWD Working Paper No. 90). [http://www.ratswd.de/download/RatSWD\\_WP\\_2009/RatSWD\\_WP\\_90.pdf](http://www.ratswd.de/download/RatSWD_WP_2009/RatSWD_WP_90.pdf) (07.02.2018); Jensen, Uwe. *Leitlinien zum Management von Forschungsdaten. Sozialwissenschaftliche Umfragedaten*. 2012. (GESIS Technical Reports 2012/07). [http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis\\_reihen/gesis\\_methodenberichte/2012/TechnicalReport\\_2012-07.pdf](http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2012/TechnicalReport_2012-07.pdf) (07.02.2018).

<sup>26</sup> So sind bspw. in der Astrophysik Fragen zu personenbezogenen Daten i. d. R. irrelevant, da keine solchen Daten erhoben werden. In den Sozialwissenschaften sollten, zumindest wenn es um Daten aus quantitativen Erhebungen geht, Fragen zu externen Erhebungsinstituten enthalten sein, da bei großen Umfragen die Datenerhebung, ggf. auch die Anonymisierung sowie erste Schritte der Datenaufbereitung oft an solche Einrichtungen ausgelagert werden.

<sup>27</sup> <https://www.rd-alliance.org/ig-domain-repositories-rda-9th-plenary-meeting> (02.02.2018).

<sup>28</sup> GitHub: <https://github.com/rdmorganiser/rdm> (02.02.2018).

ten Ebene werden diese Informationen anhand von durch die Nutzenden im jeweiligen Forschungskontext zu definierenden Projekten verwaltet. An diesen Projekten kann kollaborativ durch mehrere Personen, denen verschiedene Zugriffsrollen zugeordnet werden können, gearbeitet werden.

Die Eingabe der Informationen zu einem Projekt erfolgt in RDMO in Form eines strukturierten Interviews über das interaktive Webinterface. Das Überspringen nicht relevanter Fragen auf Basis bereits gegebener Antworten, die Nutzung verschiedener Widgets (Schieberegler, Auswahlboxen, Dropdown-Menüs usw.), die Verwendung von kontrollierten Vokabularen als Antwortoptionen usw. unterstützen eine effiziente Eingabe. Das Einfrieren der bisherigen Antworten in Datensätzen in Form von Snapshots ermöglicht eine Versionierung zu beliebigen Zeitpunkten (z. B. für Antragseinreichung, Zwischenbericht, Abschlussbericht, Publikation).

Die Ausgabe der Informationen kann über verschiedene Wege erfolgen. Die einfachste Möglichkeit ist hierbei die Ausgabe der im Interview gegebenen Antworten mit den zugehörigen Fragen als Text. Darüber hinaus können im System auch Vorlagen zu textuellen Ansichten hinterlegt werden, die aus den gegebenen Antworten neue Texte generieren, um beispielsweise mit den Anforderungen von Horizon 2020 konforme Datenmanagementpläne zu erzeugen. Es können Aufgaben definiert werden, die, durch bestimmte Antworten gesteuert, auf zukünftige Handlungsbedarfe aufmerksam machen (z. B. Abschätzung eines finanziellen Bedarfs für die Speicherung der Forschungsdaten). Schließlich werden noch programmierbare Schnittstellen (APIs) bereitgestellt, die die Nachnutzung der Daten durch andere Software-Werkzeuge ermöglichen.

Um diese Funktionalitäten zu realisieren, verfügt RDMO über ein komplexes Datenmodell<sup>29</sup>, das in Abbildung 1 dargestellt ist. Sowohl die Fragen des Fragenkataloges als auch die eingegebenen Antworten sowie Ansichten und Aufgaben referenzieren hierbei auf sogenannte Attribute und Entitäten eines baumförmigen Domänenmodells. Die Attribute dienen, wie Variablen in Programmiersprachen, als Bezeichner für einzelne Informationseinheiten (z. B. textuelle Antworten, Daten, ausgewählte Optionen). Die Entitäten fassen diese als Information zu Datensätzen zusammen. Mit diesem Datenmodell ver-

knüpft sind einerseits die in Sets organisierten Antwortoptionen und andererseits Bedingungen, die regeln, ob ein Attribut im gegebenen Kontext aktiv ist oder gegebenenfalls die verbundene Frage im Interview übersprungen werden kann. Alle im Datenmodell enthaltenen Elemente können durch ein Management-Backend editiert werden, ohne am RDMO-Quellcode Änderungen vornehmen zu müssen, wie auch Elemente hinzugefügt oder entfernt werden können. Weiter können Elemente in einem speziellen XML-Format exportiert und importiert werden, um einen Austausch zwischen den einzelnen, verteilten RDMO-Instanzen zu ermöglichen.

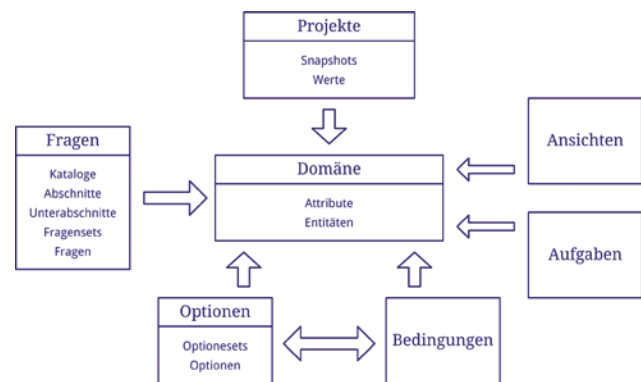


Abb. 1: RDMO-Datenmodell

Besonderer Wert wird auf die reibungslose Installation und Anpassbarkeit von RDMO durch Arbeitsgruppen und Institutionen wie Forschungsinstitute, Bibliotheken oder Hochschulen gelegt. Es ist durch sogenannte Themes möglich, das Layout sowie Farben und Schriften an das Corporate Design der betreibenden Institution anzupassen. Inhaltlich kann RDMO durch das Editieren der Elemente des Datenmodells an lokale oder fachliche Gegebenheiten angepasst werden. Authentifizierungsschnittstellen wie Shibboleth, LDAP und OAUTH2 ermöglichen die Nutzung eines lokalen Benutzermanagements für RDMO.

Für die Entwicklung von RDMO wurde auf diverse, in der Software-Community etablierte Werkzeuge und Bibliotheken zurückgegriffen. Hierzu zählen insbesondere:

- Django: ein in der Programmiersprache Python geschriebenes Framework zur Erstellung dynamischer Webseiten nach dem Model-View-Controller-Entwurfsmuster<sup>30</sup>
- Django REST Framework: eine auf Django aufbauende Bibliothek für APIs nach dem REST-Prinzip<sup>31</sup>

<sup>29</sup> Vgl. auch folgende Masterarbeit: Heger, Martin. *Datenmodellierung für Forschungsdatenmanagementpläne*. 2016. (Masterarbeit an der Fachhochschule Potsdam im Fachbereich 5 Informationswissenschaften). [http://rdmorganiser.github.io/docs/Heger\\_MA.pdf](http://rdmorganiser.github.io/docs/Heger_MA.pdf) (02.02.2018).

<sup>30</sup> <https://www.djangoproject.com> (02.02.2018).

<sup>31</sup> <http://www.django-rest-framework.org> (02.02.2018).



- AngularJS: ein JavaScript-Framework für interaktive Webseiten<sup>32</sup>
- Bootstrap: eine CSS- und Javascript-Bibliothek zur Entwicklung moderner und mobil-freundlicher Webseiten<sup>33</sup>
- Pandoc: ein Dokumentkonverter, der Exporte in viele gebräuchliche Formate ermöglicht<sup>34</sup>

Der eigentliche Entwicklungsprozess fand über die ganze Projektlaufzeit in iterativer Form und in enger Abstimmung mit der Entwicklung des Fragenkatalogs statt. Im April 2017 konnte das RDMO-Werkzeug in einer stabilen Beta-Version veröffentlicht werden. Zum Testen von RDMO wurde das Django-eigene Test-Framework verwendet und über die Dienste Travis CI<sup>35</sup> und Coveralls<sup>36</sup> ein Continuous-Integration-Workflow realisiert.

Neben der Entwicklung der Software wurde auch eine Demo-Instanz<sup>37</sup> aufgebaut. Diese Webseite wurde in der Vergangenheit (Expertengespräche, Workshops) verwendet und steht auch zukünftig weiteren Interessierten zur Verfügung. Für die Nutzung ist nur eine einfache Registrierung erforderlich. Eine Authentisierung mit ORCID<sup>38</sup> oder gängigen Social-Media-Plattformen ist ebenfalls möglich. Darüber hinaus stehen eine technische Dokumentation<sup>39</sup> und zusätzliche Skripte zur Installation mit einem Docker-Container<sup>40</sup> bereit.

### 3 Ausblick

Die erste Projektphase hatte eine Laufzeit von 18 Monaten und war personell mit insgesamt 1,33 Vollzeitäquivalenten verteilt über zwei Einrichtungen ausgestattet. Die Ergebnisse wurden regelmäßig auf Workshops und Konferenzen vorgestellt,<sup>41</sup> zuletzt auf der internationalen Tagung iPRES 2017 in Kyoto (Japan).<sup>42</sup> Auf Grund des Erfolges von RDMO

und auf Grund zahlreicher Nachfragen von interessierten Fach-Communities und Institutionen in Deutschland ist erfolgreich ein Nachfolgeantrag eingereicht worden. Die Arbeiten werden ab Ende 2017 dann mit insgesamt drei Partnern<sup>43</sup> weitergeführt.

Aufbauend auf den Ergebnissen des Vorgängerprojekts RDMO und basierend auf konkreten Rückmeldungen und Bedarfen der Zielgruppen werden im Nachfolgeprojekt diese Ziele verfolgt:

- **Erweiterung des Organisers:** Das RDMO-Werkzeug wird so erweitert, dass es Nutzenden nicht nur ermöglicht, vor dem Projektstart einen Datenmanagementplan zu entwerfen, sondern auch das Datenmanagement während des kompletten Projektverlaufs zu organisieren. Dafür ist zum einen erforderlich, die bereits anpassbaren Fragen des RDMO-Werkzeugs noch mit konkreten Aufgaben und Rollen zu verknüpfen, um einen „actionable data management plan“ zu erhalten. Zum anderen werden die zentralen Aufgaben der Kostenabschätzung, des Ingest-Prozesses und der Interoperabilität mit z. B. Nachweissystemen durch zu entwickelnde Module und ein Metadaten Application Profile besonders unterstützt. So wird RDMO nach der Beantwortung der Fragen eine Kostenabschätzung anhand der Kriterien von Datenrepositorien geben oder ein passendes Datenrepositorium anhand einer Verknüpfung mit re3data<sup>44</sup> vorschlagen. Durch die Implementierung entsprechender Schnittstellen und Identifier-Systeme wird eine Verlinkung des Datenmanagementplans mit den tatsächlichen Daten in temporären Projektspeicherorten bzw. in dauerhaften Datenrepositorien sowie eine Übernahme von Metadaten aus Forschungsinformationssystemen (FIS) ermöglicht.
- **Integration in die Infrastruktur:** Damit das RDMO-Werkzeug ein verlässliches Service-Angebot von Infrastrukturanbietern wird, wird die Integration des Werkzeugs in die jeweiligen Infrastrukturen weiter vereinfacht. Zentrale Punkte hierfür sind eine standardisierte Installation (z. B. über Docker-Container), Wartbarkeit (über einen integrierten Update-Mechanismus) und eine erweiterte Unterstützung von üblichen Authentifizierungs- und Autorisierungsverfahren.

*mentation and organisation of research data management.* iPRES in Kyoto/Japan, 25.–29. September 2017. <https://ipres2017.jp/wp-content/uploads/27Claudia-Engelhardt.pdf> (07.02.2018).

<sup>43</sup> Neu hinzugekommen ist die Bibliothek des Karlsruhe Institute of Technology (KIT).

<sup>44</sup> <https://www.re3data.org/> (02.02.2018).

<sup>32</sup> <https://angularjs.org> (02.02.2018).

<sup>33</sup> <http://getbootstrap.com> (02.02.2018).

<sup>34</sup> <http://pandoc.org> (02.02.2018).

<sup>35</sup> RDMO auf Travis CI: <https://travis-ci.org/rdmorganiser/rdmo> (02.02.2018).

<sup>36</sup> RDMO auf Coveralls: <https://coveralls.io/github/rdmorganiser/rdmo> (02.02.2018).

<sup>37</sup> <https://rdmo.aip.de/> (02.02.2018).

<sup>38</sup> <https://orcid.org/> (02.02.2018).

<sup>39</sup> <http://rdmo.readthedocs.io> (02.02.2018).

<sup>40</sup> <https://github.com/rdmorganiser/rdmo-docker> (02.02.2018).

<sup>41</sup> Vgl. RDMO-Webseite: <http://rdmorganiser.github.io/vernetzung/> (02.02.2018).

<sup>42</sup> Vgl. Engelhardt, Claudia, Harry Enke, Jochen Klar et al. *Research Data Management Organiser. A tool to support the planning, imple-*

- **Etablierung in der Community:** Sehr viele Fachdisziplinen (Biodiversität, Psychologie usw.) und Institutionen (z. B. Hochschulen, Leibniz-Einrichtungen) bzw. Wissenschaftsorganisationen (z. B. Fraunhofer Gesellschaft<sup>45</sup>) sind auf der Suche nach einem zentralen Tool zur Unterstützung von Datenmanagement und -plänen für ihre Einrichtung, die dynamisch die Dokumentation des Datenmanagements über die komplette Laufzeit von datengetriebenen Projekten abbilden und verschiedene Sichtweisen (z. B. Forschungsreferent, IT-Dienstleister) erlauben. Zudem ist es auch wichtig, international anschlussfähig zu bleiben (z. B. DMPonline<sup>46</sup> der DCC in Großbritannien, DMPTool<sup>47</sup> der California Digital Library in den USA). Des Weiteren ist es sinnvoll, die Nutzung des Tools und den konkreten Einsatz vor Ort in den Communities über Schulungen und Online-Tutorials zu fördern. Gerade der letzte Punkt als dezentraler Ansatz ist notwendig, um das RDMO-Werkzeug schnell und rasch in den tagtäglichen Einsatz zu bringen.
- **Nachhaltigkeit/Verstetigung:** Nachhaltigkeitsdiskussionen zeigen deutlich, dass daran gearbeitet werden muss, RDMO in den Communities als eigenständiges Tool soweit zu verankern, dass sich diese RDMO-Instanzen in Zukunft selbst tragen und auf diese Weise RDMO insgesamt weiterentwickelt wird. Dies wird durch die erweiterte Zusammenarbeit mit Initiativen wie E-Science Baden-Württemberg<sup>48</sup> und Einrichtungen aus den Wissenschaftsorganisationen und Hochschulen möglich. Auch gilt es, RDMO zu einer stabilen Produktreihe zu entwickeln, mit skalierbarer Anwendung und Interoperabilität zum Beispiel auch zwischen parallelen Instanzen innerhalb einer Disziplin.

Insgesamt gilt es im Nachfolgeprojekt einerseits, eine dezentrale Struktur aufzubauen, die die technische und organisatorische Weiterentwicklung des RDMO-Werkzeugs garantiert, und andererseits eine aktive Community zu etablieren, die alle relevanten Aspekte des aktiven Forschungsdatenmanagements und deren Stakeholder berücksichtigt. Im besten Fall entsteht so in Deutschland, quer über die Fachdisziplinen und Wissenschaftsorganisationen hinweg, ein gemeinsamer Prozess, der basierend auf transparenten Entscheidungen Forschungsdatenmanagement und damit auch die Reproduzierbarkeit,

<sup>45</sup> Hierzu liegt ein aktueller Letter of Intent vor.

<sup>46</sup> <https://dmponline.dcc.ac.uk/> (02.02.2018).

<sup>47</sup> <https://dmptool.org/> (02.02.2018).

<sup>48</sup> <https://mwk.baden-wuerttemberg.de/de/forschung/forschungslandschaft/e-science/> (02.02.2018).

Validierbarkeit und Nachnutzung von Forschungsdaten erlaubt. Damit würde auch von den Forschenden und den Infrastruktureinrichtungen in Deutschland ein wesentlicher Beitrag zu den Diskussionen und Initiativen um Open Science<sup>49</sup> geleistet.

## Autoreninformationen



**Prof. Dr. rer. nat. Heike Neuroth**

Fachhochschule Potsdam (FHP)  
 Fachbereich Informationswissenschaften  
 Kiepenheuer Allee 5  
 14469 Potsdam  
[neuroth@fh-potsdam.de](mailto:neuroth@fh-potsdam.de)  
<http://orcid.org/0000-0002-3637-3154>



**Claudia Engelhardt**

Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB)  
 Abteilung Forschung und Entwicklung  
 Platz der Göttinger Sieben 1  
 37073 Göttingen  
[claudia.engelhardt@sub.uni-goettingen.de](mailto:claudia.engelhardt@sub.uni-goettingen.de)  
<http://orcid.org/0000-0002-3391-7638>



**Dr. Jochen Klar**

Leibniz-Institut für Astrophysik Potsdam (AIP)  
 Supercomputing und E-Science  
 An der Sternwarte 16  
 14482 Potsdam  
[jklar@aip.de](mailto:jklar@aip.de)  
<http://orcid.org/0000-0002-5883-4273>



**Jens Ludwig**

Staatsbibliothek zu Berlin (SBB)  
 Haus Potsdamer Straße  
 Informations- und Datenmanagement  
 Potsdamer Str. 33  
 10785 Berlin  
[jens.ludwig@sbb.spk-berlin.de](mailto:jens.ludwig@sbb.spk-berlin.de)  
<http://orcid.org/0000-0002-2998-1696>



**Dr. Harry Enke**

Leibniz-Institut für Astrophysik Potsdam (AIP)  
 Supercomputing und E-Science  
 An der Sternwarte 16  
 14482 Potsdam  
[henke@aip.de](mailto:henke@aip.de)  
<http://orcid.org/0000-0002-2366-8316>

<sup>49</sup> Vgl. z. B. <http://ec.europa.eu/programmes/horizon2020/en/h2020-section/open-science-open-access> (02.02.2018).