

Ein kooperativ nutzbares Werkzeug zur Erstellung von Datenmanagementplänen

Franziska Ackermann¹, Dennis Wehrle²

¹Kommunikations- und Informationszentrum (kiz), Universität Ulm

²Professur für Kommunikationssysteme, Albert-Ludwigs-Universität Freiburg

Zusammenfassung

Wissenschaftler müssen sich zunehmend mit Richtlinien und Empfehlungen zum Datenmanagement auseinandersetzen, die von Drittmittelgebern, ihren Forschungseinrichtungen, von Fachdisziplinen, Verlagen und wichtigen Vertretern aus Wissenschaft und Politik ausgehen. Ein prominentes Beispiel ist das neue Element des Datenmanagementplans (DMP) in „Horizon 2020“, der auf verschiedene Aspekte wie z.B. Beschreibung, Backup und Nachnutzung der Daten eingeht. Projekte, die am „Open Data Pilot“ der Europäischen Kommission teilnehmen, müssen ein solches Dokument über das geplante Datenmanagement im Projekt einreichen und regelmäßig aktualisieren. Datenmanagementpläne haben sich im Vereinigten Königreich und den USA bereits zu einem Standard entwickelt. Dort können Wissenschaftler auf Software-Werkzeuge zurückgreifen (sog. DMP-Werkzeuge), die Vorlagen und Hilfestellung beim Erstellen eines passenden Datenmanagementplans beinhalten. Das vom Land Baden-Württemberg geförderte Projekt „Landesweit koordinierte Strukturen für Nachweis und effiziente Nachnutzung von Forschungsdaten“ (01/2014 – 02/2016) zwischen der Universität Ulm und der Universität Freiburg beinhaltet neben andere Zielen auch die Suche nach einer Lösung für ein DMP-Werkzeug, das von mehreren Einrichtungen in Baden-Württemberg kooperativ genutzt werden kann. Nach einem Vergleich bereits vorhandener Entwicklungen entschied man sich im Projekt dazu, den Quellcode des „DMPTool“ der California Digital Library nachzunutzen, das in den USA von einer Vielzahl von Einrichtungen gemeinsam genutzt wird. Die Software wurde von den Projektpartnern Ulm und Freiburg technisch und inhaltlich angepasst und beinhaltet z.B. eine Vorlage für „Horizon 2020“, die Wissenschaftler bei der Erstellung eines DMPs für ihr Projekt unterstützen kann. Über die Authentifizierung mittels Shibboleth/bwIDM soll das DMP-Werkzeug von Wissenschaftlern der teilnehmenden Einrichtungen unkompliziert für die Erstellung von Datenmanagementplänen genutzt werden können.

Abstract

Researchers increasingly face data management policies from funders, their research institutions, research communities, publishers and players from science and politics. A prominent example is the new element of a data management plan (DMP) in “Horizon 2020”, which includes aspects such as description, backup and re-use of data. Projects participating in the “Open Data Pilot” of the European Commission are asked to deliver such a document outlining their plan for data management in the project, and to keep it updated. Data management plans have become a standard in the UK and the USA, where researchers are offered tools that provide templates and guidance for writing a suitable data management plan. Among other goals, Ulm University and the University of Freiburg aimed to find a solution for a DMP tool that can be used cooperatively by several institutions in Baden-Württemberg in the project “Landesweit koordinierte Strukturen für Nachweis und effiziente Nachnutzung von Forschungsdaten”¹ (01/2014 – 02/2016), which was funded by the state Baden-Württemberg. After a comparison of existing DMP tools, the project partners decided to use the source code from “DMPTool” from the California Digital Library, which is used cooperatively by a large number of institutions in the USA. The project partners Ulm and Freiburg adapted “DMPTool” technically and content-wise. The adapted tool includes a template for “Horizon 2020”, for instance, which can support researchers when creating a data management plan for their project. It is intended to provide researchers from participating institutions with a comfortable way to use the tool for creating data management plans via Shibboleth/bwIDM authentication.

1. Motivation

Immer häufiger werden Wissenschaftler aufgefordert darzulegen, wie sie mit den erhobenen Daten in ihrem Forschungsvorhaben umgehen werden und ob und unter welchen Bedingungen die Daten nach Projektende für die Öffentlichkeit nachnutzbar sind. Nach der Praxis beim Datenmanagement fragen beispielsweise Drittmittelgeber wie die Deutsche Forschungsgemeinschaft (DFG) und die Europäische Union (EU).² Desweiteren verabschieden Forschungseinrichtungen in den letzten Jahren zunehmend Richtlinien und Empfehlungen zum Datenmanagement (Data Policies), in denen häufig die Erstellung eines Datenmanagementplans (DMP) gefordert wird.

¹ Statewide development of coordinated structures for indexing and re-use of research data

² DFG: http://www.dfg.de/formulare/54_01/54_01_de.pdf, Zugriff 25.04.2016

EU: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf, Zugriff 25.04.2016

Dabei handelt es sich um ein Dokument, das beschreibt, welche Art von Daten im Forschungsvorhaben genutzt und erhoben werden, wie sie gesichert werden, welche ethischen oder rechtlichen Aspekte eine Rolle spielen, wie die Nachnutzung der Daten geplant ist, etc. Neben Drittmittelgebern und Forschungseinrichtungen gibt es zudem einzelne Zeitschriften mit einer Data Policy, die Richtlinien und Empfehlungen für die Autoren zum Umgang mit den Daten enthält, die zu einer eingereichten Publikation gehören.³ Daneben gibt es weitere Empfehlungen zum Umgang mit Forschungsdaten, beispielsweise einzelner Fachdisziplinen oder von wichtigen Vertretern aus Wissenschaft und Forschung, wie der DFG⁴ und der Hochschulrektorenkonferenz (HRK)⁵.

Hinter der Forderung von Drittmittelgebern wie der EU und der DFG nach einem professionellen Datenmanagement und „Open Data“ steckt die Idee, dass öffentlich geförderte Forschung nach Möglichkeit wiederum öffentlich verfügbar und nachnutzbar sein soll. Wenn Daten stärker als bisher für eine Nachnutzung verfügbar wären, würde dies beispielsweise dazu beitragen, dass publizierte Ergebnisse besser nachvollzogen und verifiziert, dass kostspielige Neuerhebungen von Daten vermieden und einmalige Daten (z.B. Wetterdaten) mehrfach ausgewertet werden könnten.

1.1 Data Policies von Drittmittelgebern

Im aktuellen Rahmenprogramm für Forschung und Innovation „Horizon 2020“ der Europäischen Kommission gibt es als neues Element den Datenmanagementplan. Von Projekten, die dort am „Open Data Pilot“ teilnehmen, wird erwartet, dass sie spätestens ein halbes Jahr nach Projektbeginn einen Datenmanagementplan vorlegen und diesen aktuell halten. Die Europäische Kommission bietet einen Leitfaden zum Datenmanagement an, der unter anderem konkrete Fragen zum Umgang mit Forschungsdaten enthält, die von den Wissenschaftlern beantwortet werden sollen.⁶ Dieser Pilot der Europäischen Kommission zu offenen Daten ist insofern interessant, als dass der vorausgegangene Pilot zu Open Access für Publikationen inzwischen in einen Standard für sämtliche bewilligte Projekte überführt wurde.

³ Public Library of Science (PLoS), American Geophysical Unit (AGU) u.a.

⁴ http://www.dfg.de/foerderung/antragstellung_begutachtung_entscheidung/antragstellende/antragstellung/nachnutzung_forschungsdaten/, Zugriff 25.04.2016

⁵ http://www.hrk.de/uploads/tx_szconvention/Empfehlung_Forschungsdatenmanagement_final_Stand_11.11.2015.pdf, Zugriff 25.04.2016

⁶ Leitfaden: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf, Version 2.1 vom 15. Februar 2016, Zugriff 25.04.2016

Eine ähnliche Formulierung wie die der Europäischen Kommission findet sich im Leitfaden für die Antragsstellung der Deutschen Forschungsgemeinschaft (DFG):

„Wenn aus Projektmitteln systematisch Forschungsdaten oder Informationen gewonnen werden, die für die Nachnutzung durch andere Wissenschaftlerinnen und Wissenschaftler geeignet sind, legen Sie bitte dar, ob und auf welche Weise diese für andere zur Verfügung gestellt werden. Bitte berücksichtigen Sie dabei auch - sofern vorhanden - die in Ihrer Fachdisziplin existierenden Standards und die Angebote existierender Datenrepositorien oder Archive“ (DFG 2014, S.5)

Die Erwartungen der DFG sind detaillierter in den „Leitlinien zum Umgang mit Forschungsdaten“ formuliert, die im September 2015 vom Senat der DFG verabschiedet wurden.⁷ Diese Leitlinien zeigen zudem auf, wie die DFG die Wissenschaftler beim Datenmanagement unterstützt und beinhalten einen Appell an die wissenschaftlichen Fachgemeinschaften, disziplinspezifische Regularien zu entwickeln. Einzelne Fachdisziplinen sind dieser Aufforderung bereits nachgekommen, z.B. die Biodiversität, die Erziehungswissenschaften und die Digital Humanities (Digitale Editionen, Sprachkorpora).⁸

1.2 Data Policies an Forschungseinrichtungen in Deutschland

Neben Drittmittelgebern erwarten auch Forschungseinrichtungen verstärkt einen Datenmanagementplan oder ein vergleichbares Dokument. In Deutschland haben verschiedene Universitäten Richtlinien zum Datenmanagement verabschiedet (s. Tabelle). Data Policies gibt es zudem an weiteren, nicht-universitären Forschungseinrichtungen, wie z.B. dem Deutschen GeoForschungsZentrum in Potsdam und dem Robert Koch-Institut in Berlin.

⁷ http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf, Zugriff 25.04.2016

⁸ Der Appell an die wissenschaftlichen Fachgemeinschaften findet sich in den Richtlinien zum Forschungsdatenmanagement und zusätzlich auf der Webseite: http://www.dfg.de/foerderung/antragstellung_begutachtung_entscheidung/antragstellende/antragstellung/nachnutzung_forschungsdaten/, Zugriff 10.02.2016. Die bereits erarbeiteten disziplinspezifischen Regularien sind dort ebenfalls verlinkt.

Der Großteil der Universitäten mit einer Data Policy fordert, dass ein Datenmanagementplan erstellt wird:

Universität Bielefeld	„Ein „Data Management Plan“ und fachspezifische Richtlinien** für Einrichtungen und Projekte soll durch die jeweils verantwortliche Leitung bereitgestellt werden – vor allem in jedem neuen, Daten-intensiven Drittmittelantrag.“ (Universität Bielefeld 2011)
Universität Göttingen	„Forschungsprojekte mit Forschungsdaten erfordern einen Datenmanagementplan, der u.a. die Zugangsrechte und -vorbehalte der Forschungsdaten darlegt.“ (Universität Göttingen 2014)
Universität Heidelberg	„Teil jedes Forschungsprojektes ist ein Plan für das Datenmanagement, der explizit adressiert, wie die Akkuratheit, Vollständigkeit, Authentizität, Integrität, Vertraulichkeit, Veröffentlichung und der offene Zugang von Daten gehandhabt werden. Dabei werden fachspezifische Besonderheiten berücksichtigt.“ (Universität Heidelberg 2014)
Universität Kiel	„Das Forschungsdatenmanagement wird nach den aktuellen fachlichen Standards unter Einhaltung der Empfehlungen der DFG zur Sicherung der guten wissenschaftlichen Praxis durchgeführt und in einem Datenmanagement-Plan dokumentiert. Die Verantwortung für das Forschungsdatenmanagement liegt bei der Leitung der jeweiligen Forschungsaktivität.“ (Universität Kiel 2015)
TU Darmstadt	„Die Universität empfiehlt die Aufstellung eines Plans für den Umgang mit Forschungsdaten bereits vor Beginn eines Forschungsvorhabens, um einen systematischen und nachhaltigen Umgang mit den Daten zu gewährleisten. Ein Forschungsdatenmanagementplan enthält eine Beschreibung aller relevanten im Laufe des Vorhabens entstehenden Daten sowie ein Konzept für den Umgang mit ihnen im Hinblick auf Genauigkeit, Vollständigkeit, Authentizität, Integrität, Vertraulichkeit und Verbleib, sowie Urheber- und Nutzungsrechte. Dabei sind fachspezifische Besonderheiten und Standards zu berücksichtigen und der Plan dem jeweils aktuellen Verlauf des Vorhabens anzupassen.“ (TU Darmstadt 2015)

Es bleibt festzuhalten, dass Wissenschaftler zunehmend mit Richtlinien und Empfehlungen zum Datenmanagement konfrontiert werden, die für sie beispielsweise als Angehörige einer Forschungseinrichtung, als Antragsteller eines Drittmittelprojekts, oder als Angehörige einer Fachdisziplin gelten. Für die Erstellung von Datenmanagementplänen gibt es Werkzeuge, welche diese Bedingungen abbilden können und Wissenschaftler durch Beispieltex-te, institutionsspezifische Hinweise, weiterführenden Links und zusätzliche Hilfestellung bei der Erstellung eines passenden DMPs unterstützen können.

2. Data Policies im anglo-amerikanischen Raum

Im anglo-amerikanischen Raum ist es bereits stark verankert, dass im Rahmen von Drittmittelprojekten Datenmanagementpläne erstellt werden. Die National Science Foundation ist hierfür ein prominentes Beispiel.⁹ Weitere Drittmittelgeber mit Vorgaben zu einem Datenmanagementplan sind die Alfred P. Sloan Foundation, das Department of Education, Department of Energy, und andere.¹⁰ In den USA und dem Vereinigten Königreich gibt es bereits etablierte Lösungen, die den Wissenschaftlern die Erstellung eines Datenmanagementplans erleichtern, sogenannte *DMP-Werkzeuge*. Sie nehmen den Projektteilnehmern nicht ab, sich mit den Themen des Datenmanagements auseinanderzusetzen, bieten jedoch verschiedene Vorlagen und zusätzliche Hilfestellungen in Form von Erläuterungen, Links, Beispiel- und Musterantworten an. Derzeit gibt es zwei etablierte DMP-Werkzeuge, die in den USA bzw. im Vereinigten Königreich kooperativ von jeweils mehr als 180 Einrichtungen genutzt werden. Diese sind das „DMPTool“¹¹ der California Digital Library (USA) und „DMPonline“¹² des Digital Curation Centers (UK). Die dort angebotenen Vorlagen und Hilfestellungen können von den teilnehmenden Einrichtungen jeweils an die eigene Situation angepasst werden, sodass die Wissenschaftler vor Ort die benötigte Unterstützung erhalten können.

⁹ <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>, Zugriff 18.02.2016

¹⁰ Alfred P. Sloan Foundation: <http://www.sloan.org/apply-for-grants/grant-proposals/>
Department of Education: http://ies.ed.gov/funding/datasharing_implementation.asp
Department of Energy: <http://www.energy.gov/datamanagement/doe-policy-digital-research-data-management>, Zugriff 26.04.2016

¹¹ <https://dmptool.org/>

¹² <https://dmponline.dcc.ac.uk/>

Das vom Land Baden-Württemberg geförderte Projekt „Landesweit koordinierte Strukturen für Nachweis und effiziente Nachnutzung von Forschungsdaten“ (01.01.2014 – 29.02.2016) beinhaltet als Teilaspekt die Suche nach einer Lösung für ein DMP-Werkzeug zur Erstellung von Datenmanagementplänen, das von mehreren Einrichtungen im Land Baden-Württemberg kooperativ genutzt werden kann.

3. Anforderungen an ein kooperatives DMP-Werkzeug

Ein wesentlicher Aspekt bei der Suche nach einem DMP-Werkzeug war der kooperative Charakter. Im Projekt wurde explizit nach einer existierenden Softwarelösung gesucht, die nur an einem Ort installiert und gepflegt werden muss und dennoch von einer Vielzahl von Einrichtungen genutzt werden kann. Dadurch muss nicht jede Einrichtung eine lokale Installation pflegen. Vielmehr können Ressourcen für die Bereitstellung und Weiterentwicklung des DMP-Werkzeugs gebündelt werden. Gleichzeitig sollte die Softwarelösung so flexibel sein, dass inhaltliche Anpassungen durch die teilnehmenden Einrichtungen möglich sind. So soll eine Einrichtung beispielsweise selbst Vorlagen für Datenmanagementpläne erstellen, anpassen und Hilfestellungen (Erläuterungen, Links, Beispiel- und Musterantworten) einbinden können, die auf die lokalen Gegebenheiten ausgerichtet sind.

Aus Sicht der Nachhaltigkeit ist eine möglichst große Community, welche das DMP-Werkzeug nutzt und unterstützt, von Vorteil. Eine Open Source Lösung ist ebenfalls von Vorteil, da hier zum einen durch die transparente Funktionsweise der Software bei Bedarf Anpassungen vorgenommen werden können und zum anderen einer Obsoleszenz entgegengewirkt werden kann.

Ein weiterer Aspekt, den es zu berücksichtigen gilt, betrifft die Datensicherheit und den Datenschutz. Die Universität Bielefeld berichtete, dass Wissenschaftler bei der Nutzung ihres DMP-Werkzeugs besonderen Wert auf diese beiden Punkte legten.¹³ Dies ist nachvollziehbar, da die Datenmanagementpläne Informationen zu den Daten im Projekt und den Beteiligten enthalten, die wiederum Rückschlüsse auf das Projektvorhaben ermöglichen.

¹³ Dieser Punkt findet sich im Fazit des Vortrags, der im Rahmen des Treffens der DINI/nestor AG Forschungsdaten im März 2015 von Johanna Vompras gehalten wurde. Folien: <http://www.forschungsdaten.org/images/3/31/04--vompras--lessons-learnt-bielefeld.pdf>, Zugriff 10.02.2016

Damit das deutsche Datenschutzrecht Anwendung findet, wurde im Projekt die Prämisse gesetzt, dass die beiden Projektteilnehmer Ulm und Freiburg für das Testsystem eine eigene Installation an einem der beiden Standorte vornehmen, welche kooperativ genutzt werden kann. In Zukunft kann eine Migration auf einen anderen Server erfolgen, sofern dies keine Probleme bezüglich des Datenschutzes mit sich bringen würde.

Um Wissenschaftlern eine einfache Nutzung zu ermöglichen, wurde eine komfortable Authentifizierungsmethode angestrebt. Über bwIDM bzw. Shibboleth können sich Wissenschaftler der Universitäten in Baden-Württemberg beispielsweise mit ihren gewohnten institutionellen Login-Daten anmelden und müssen keinen separaten Account mit Passwort anlegen, was einen unkomplizierten Einstieg zur Erstellung eines Datenmanagementplans ermöglichen würde. Die Menüführung und Funktionen sollten ebenfalls weitestgehend selbsterklärend und schnell erlernbar sein.

4. Vergleich möglicher Kandidaten für ein kooperatives DMP-Werkzeug

Es gibt etablierte DMP-Werkzeuge, die bereits kooperativ genutzt werden und weitere Lösungen, die lokal an einer bestimmten Einrichtung zum Einsatz kommen. Bei den etablierten, kooperativ genutzten Angeboten handelt es sich wie eingangs erwähnt zum einen um „DMPonline“, dem DMP-Werkzeug des Digital Curation Center (UK) für Einrichtungen aus dem Vereinten Königreich, und zum anderen um das „DMPTool“, das von der California Digital Library (USA) für eine Vielzahl amerikanischer Universitäten und Einrichtungen bereitgestellt wird. Bei beiden DMP-Werkzeugen ist der Quellcode nachnutzbar, wovon teilweise bereits gebraucht gemacht wurde.¹⁴ In Deutschland gibt es an lokal eingesetzten DMP-Werkzeugen zum Beispiel das selbstentwickelte, in die Infrastruktur vor Ort eingebundene DMP-Werkzeug „TUB-DMP“ an der Technischen Universität Berlin, dessen Quellcode frei verfügbar ist.¹⁵ An der Universität Bielefeld kommt ebenfalls ein lokales DMP-Werkzeug zum Einsatz.¹⁶ Neben entsprechenden Softwareangeboten einzelner Einrichtungen gibt es auch den Ansatz, ein auf die Fachdisziplin zugeschnittenes DMP-Werkzeug anzubieten, wie es beispielsweise bei CLARIN-D der Fall ist.¹⁷

¹⁴ So gibt es z.B. eine spanische Version von DMPonline unter <http://dmp.consorciomadrono.es/>, Zugriff 21.04.2016

¹⁵ <https://github.com/tuub/TUB-DMP>, Zugriff 11.02.2016

¹⁶ <https://data.uni-bielefeld.de/de/data-management-plan>, Zugriff 11.02.2016

¹⁷ <http://www.clarin-d.de/de/aufbereiten/datenmanagementplan-entwickeln>, Zugriff 21.04.2016

Aufgrund der im vorausgegangenen Abschnitt genannten Kriterien an ein DMP-Werkzeug, insbesondere des kooperativen Charakters und des großen Nutzerkreises, rückten die beiden etablierten Angebote „DMPonline“ und „DMPTool“ in den Fokus und wurden genauer untersucht. Sie wurden zum Zeitpunkt des Vergleichs von jeweils weit über hundert Einrichtungen kooperativ genutzt – inzwischen sind es jeweils über 180 Einrichtungen – und ähneln sich in ihren Funktionen stark. Beide Softwarelösungen ermöglichen die Erstellung von weiteren Vorlagen durch die teilnehmenden Einrichtungen, sodass hier gezielt auf bestimmte Forschungsförderer und Fachdisziplinen eingegangen werden kann. „DMPTool“ bot jedoch einige interessante Features, die bei „DMPonline“ nicht zu finden waren, z.B. die Möglichkeit, einen Plan zu duplizieren, was für die Überarbeitung von Plänen hilfreich ist, oder die Option, einen Plan durch institutionelle Ansprechpartner begutachten zu lassen (Review-Prozess). Aufgrund der Möglichkeit die Software auf Basis der vorhandenen MIT-Lizenz den Bedürfnissen entsprechend anpassen zu können sowie des Vergleichs anhand der aufgestellten Kriterien wurde im Projekt entschieden, den Quellcode des „DMPTool“ der California Digital Library nachzunutzen und das DMP-Werkzeug technisch und inhaltlich anzupassen, damit es der Situation im deutschsprachigen Raum gerecht wird und von den Projektpartnern kooperativ genutzt werden kann. Für das im Projekt angepasste DMP-Werkzeug verwenden wir an dieser Stelle die Bezeichnung „DMP-Tool“, um es von den verschiedenen DMP-Werkzeugen und dem originalen „DMPTool“ begrifflich abzugrenzen.

Inzwischen haben sich die Datenmodelle der beiden DMP-Werkzeuge „DMPonline“ und „DMPTool“ stark aneinander angenähert, sodass das Digital Curation Center und die California Digital Library die Absicht erklärt haben, künftig gemeinsam an einer Lösung zu arbeiten. Es soll ein Prototyp auf Basis des Quellcodes von „DMPonline“ entstehen, der um Funktionalitäten von „DMPTool“ ergänzt wird, wie beispielsweise die Review-Funktion.¹⁸

¹⁸ Simms et al. *The Future of Data Management Planning: Tools, Policies, and Players*. International Digital Curation Conference (IDCC) 2016, Practice Paper, <https://dmptool.files.wordpress.com/2016/02/idcc16-simms-et-al-29-jan1.pdf>, S.7.

5. Was kann das DMP-Tool leisten?

Die folgende Beschreibung bezieht sich sowohl auf das „DMPTool“ der California Digital Library als auch auf die im Projekt vorgenommene Anpassung. Das DMP-Tool ist webbasiert und kann über eine einzige zentrale Installation von einer Vielzahl von Einrichtungen gemeinsam genutzt und den Wissenschaftlern zur Erstellung von Datenmanagementplänen angeboten werden. Der kooperative Charakter des DMP-Tools zeigt sich nicht nur in der technischen Komponente, sondern auch in der inhaltlichen. Teilnehmende Einrichtungen können bestehende DMP-Vorlagen nachnutzen und anpassen. Sie können Musterpläne erstellen und diese im DMP-Tool als Orientierung bereitstellen. Wissenschaftler wiederum können ihre Pläne als Beispiele auf der Webseite anbieten. Die Sichtbarkeit kann dabei auf die Wissenschaftler der eigenen Einrichtung beschränkt werden, komplett öffentlich oder komplett privat (Mitbesitzer mit eingeschlossen) sein.

Wissenschaftler können sich am DMP-Tool bequem mit ihren institutionellen Login-Daten anmelden. In der Weboberfläche können sie aus verschiedenen DMP-Vorlagen auswählen, z.B. zugeschnitten auf einen Forschungsförderer oder eine Fachdisziplin. Durch zusätzliche Hilfestellungen in Form von Links, Erläuterungen, Muster- und Beispielantworten werden die Nutzer bei der Erstellung des Plans unterstützt. Es können mehrere Personen an einem Plan schreiben, wobei eine Person der Hauptbesitzer ist. Diese kann weitere Personen, die bereits einen DMP-Tool Account¹⁹ besitzen müssen, als Mitbesitzer des Plans hinzufügen. Der Plan wird in der Applikation gespeichert und ist dadurch zentral abgelegt und für alle berechtigten Personen im Web einsehbar und veränderbar.

Es ist möglich, über das DMP-Tool bei Ansprechpartnern der Einrichtung Feedback zu den erstellten Plänen einzuholen. Diese Feedback-Funktion wird auf Ebene der Vorlagen festgelegt und bietet drei Möglichkeiten, die für die Pläne gelten, die mit der Vorlage erstellt wurden: kein Feedback, optionales Feedback oder obligatorisches Feedback. Hierdurch können die Chancen auf einen erfolgreichen Projektantrag und –verlauf erhöht werden. Datenmanagementpläne können als PDF oder DOCX exportiert werden.

¹⁹ Der Account wird beim erstmaligen Login angelegt.

Das DMP-Tool arbeitet mit einem differenzierten Rollenmodell. Neben einem oder mehreren *Administrator(en)* für die gesamte Webapplikation gibt es auch auf der Ebene der einzelnen Einrichtungen *institutionelle Administratoren*, die im DMP-Tool für die eigene Einrichtung die umfassendsten Berechtigungen haben. Wissenschaftler nehmen wie oben erwähnt die Rolle des Hauptbesizers oder Mitbesizers eines Plans ein. Feedback geben können sogenannte *Reviewer*. Reviewer erhalten die Pläne zur Begutachtung, welche von Wissenschaftlern der eigenen Einrichtung mit Vorlagen erstellt wurden, die die Review-Option anbieten. Die Erstellung der Vorlagen wird vom *Template Editor* vorgenommen, während die zusätzlichen Hilfestellungen (Links, Beispiel- und Musterantworten, Erläuterungen) vom *Resource Editor* erstellt werden. Eine Person kann selbstverständlich mehrere Rollen einnehmen.

6. Praktische Umsetzung und prototypischer Betrieb

Das von der California Digital Library angebotene „DMPTool“ wurde technisch und inhaltlich angepasst. Für die Anpassungen wurde ein *fork*²⁰ auf GitHub²¹ erstellt, in dem die Änderungen transparent nachvollziehbar sind. Von Anfang an fand eine Trennung bei der Entwicklung zwischen den optischen und den funktionalen Änderungen innerhalb von sogenannten *Branches* statt. Ziel dieser Trennung war es, dass die Erweiterung des Codes um die Mehrsprachigkeit unabhängig von etwaigen optischen Änderungen gepflegt wird, damit die funktionalen Änderungen so generisch wie möglich gehalten werden und somit für die Community von Interesse ist. Am Ende können die im Projekt vorgenommenen Anpassungen durch die Entwickler des ursprünglichen DMP-Werkzeugs in den Branch übernommen werden, der dem aktuellen Stand des von der California Digital Library verwendeten „DMPTool“ entspricht. Dies würde in Zukunft parallele Arbeiten am Code verhindern und bewirken, dass beide Installationen – die amerikanische und die im Projekt angepasste Softwarelösung – von Updates profitieren können. Für das angepasste DMP-Tool würde dies bedeuten, dass Neuerungen aus dem „DMPTool“ leicht übernommen werden können, da lediglich eine Integration der optischen Änderungen in die neue Version erfolgen muss.

²⁰ Ein fork ist eine Abspaltung und somit eine Kopie einer Software, die ab diesem Zeitpunkt unabhängig von der originalen Software entwickelt wird.

²¹ DMP-Tool, <https://github.com/derDaywalker/dmptool/>, Zugriff 22.04.2016

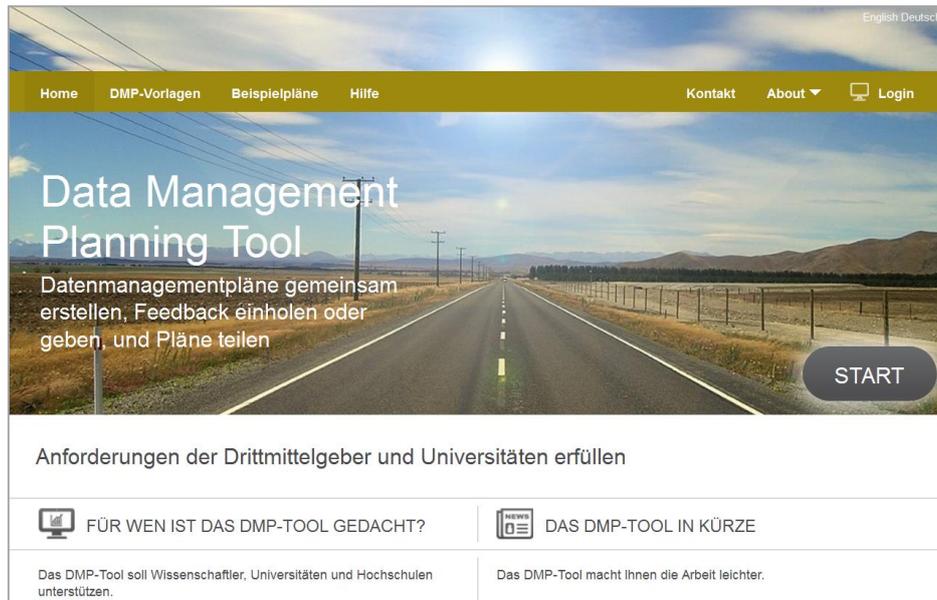


Abb.1: Startseite des DMP-Tools

Das DMP-Tool wurde im Rechenzentrum der Universität Freiburg installiert und wird dort technisch betreut. Im prototypischen Testsystem wird das DMP-Tool aktuell von den beiden Projektpartnern erprobt, der Universität Ulm und der Universität Freiburg. Statt dem Shibboleth-Login für amerikanische Einrichtungen wurde bwIDM angebunden, die Shibboleth-Förderung des Landes Baden-Württemberg. Aus Datenschutzgründen wurde eine zusätzliche Seite erstellt und so eingebunden, dass beim erstmaligen Login um Zustimmung zu den Hinweisen zum Datenschutz gebeten wird. Auch die Startseite wurde leicht modifiziert und das Menü verschlankt.

Ein wesentlicher Punkt war die Realisierung einer zweisprachigen Oberfläche (Englisch und Deutsch). Hierzu musste das DMP-Tool zunächst internationalisiert werden, d.h. die Textfragmente im Quellcode wurden herausgesucht und durch sogenannte Internationalisierungs-Variablen ersetzt. Die Texte wurde ins Deutsche übersetzt und – wo sinnvoll - auch inhaltlich angepasst. Insgesamt wurden mehr als 1050 Textstellen mit mehr als 5100 Wörtern sowie zehn komplett statische Seiten (Hilfeseiten, Datenschutzerklärung, ...) übersetzt. Durch die im Projekt vorgenommene Internationalisierung können nun auf leichte Weise weitere Sprachen hinzugefügt werden.

7. DMP-Vorlagen („Horizon 2020“ u.a.)

Für die Zielgruppe des DMP-Tools sind wie eingangs beschrieben verschiedene Richtlinien und Empfehlungen zum Datenmanagement relevant. Im Projekt wurden Vorlagen für „Horizon 2020“ erstellt, welche die Fragen und Hinweise aus dem Leitfaden der Europäischen Kommission zum Datenmanagement enthalten. Zusätzlich wurde eine generische Vorlage in das DMP-Tool eingepflegt, die für die Erstellung von Datenmanagementplänen für Projekte genutzt werden kann, bei denen keine konkreten Vorgaben durch Drittmittelgeber oder von anderer Seite existieren. Die Erstellung weiterer Vorlagen ist ein Teilziel von Projekten aus der E-Science-Ausschreibung des Landes Baden-Württemberg. Dort sollen beispielsweise in Zusammenarbeit mit Wissenschaftlern aus der Computerlinguistik (Replay-DH) und den Ingenieurwissenschaften/ der Informatik (bwFDM-Soft) jeweils geeignete Vorlagen erstellt werden, welche sich an den in den Projekten zu entwickelten Archivierungsabläufen orientieren. Ideal ist es, wenn die jeweiligen Fachdisziplinen und die Drittmittelgeber bzw. die Autoren der Data Policies in die Erstellung der Vorlagen einbezogen werden. Wenn es bereits eine detaillierte Anleitung für einen Datenmanagementplan gibt, wie es z.B. bei „Horizon 2020“ der Fall ist, können die entsprechenden Fragen und Hinweise des Forschungsförderers für die Erstellung einer Vorlage verwendet werden, die dann speziell von Projekten dieses Forschungsförderers genutzt werden kann. Die auf diese Weise erstellte Vorlage kann wiederum von jeder teilnehmenden Einrichtung angepasst werden, indem z.B. institutionsspezifische Hinweisen ergänzt werden.

Die im DMP-Tool angebotenen Vorlagen zu „Horizon 2020“ können durch jede teilnehmende Einrichtung um einrichtungsspezifische Hinweise wie Erläuterungen, Links, Muster- und Beispielantworten ergänzt werden. Gibt es ein Datenrepositorium an der Universität, kann hierfür beispielsweise ein Mustertext verfasst werden, den die Wissenschaftler der Einrichtung für ihren Plan nutzen können. Neben einrichtungsspezifischen Hinweisen bietet das kooperative DMP-Tool auch die Möglichkeit, Hilfestellungen zu erstellen, die global für alle Einrichtungen sichtbar sind. Hiervon wurde bei der Vorlage für „Horizon 2020“-Projekte Gebrauch gemacht, damit allgemeingültige Hilfestellungen nicht von jeder Einrichtung separat eingepflegt werden müssen.

Die Europäische Kommission sieht vor, dass Projekte im „Open Data Pilot“ den Datenmanagementplan aktuell halten und zu bestimmten Zeitpunkten erneut abgeben.

Das DMP-Tool ermöglicht ein leichtes Duplizieren von Plänen, sodass die ursprüngliche Version beibehalten werden kann. Die duplizierte Version beinhaltet bereits die bisherigen Antworten und kann so auf unkomplizierte Weise aktualisiert werden.

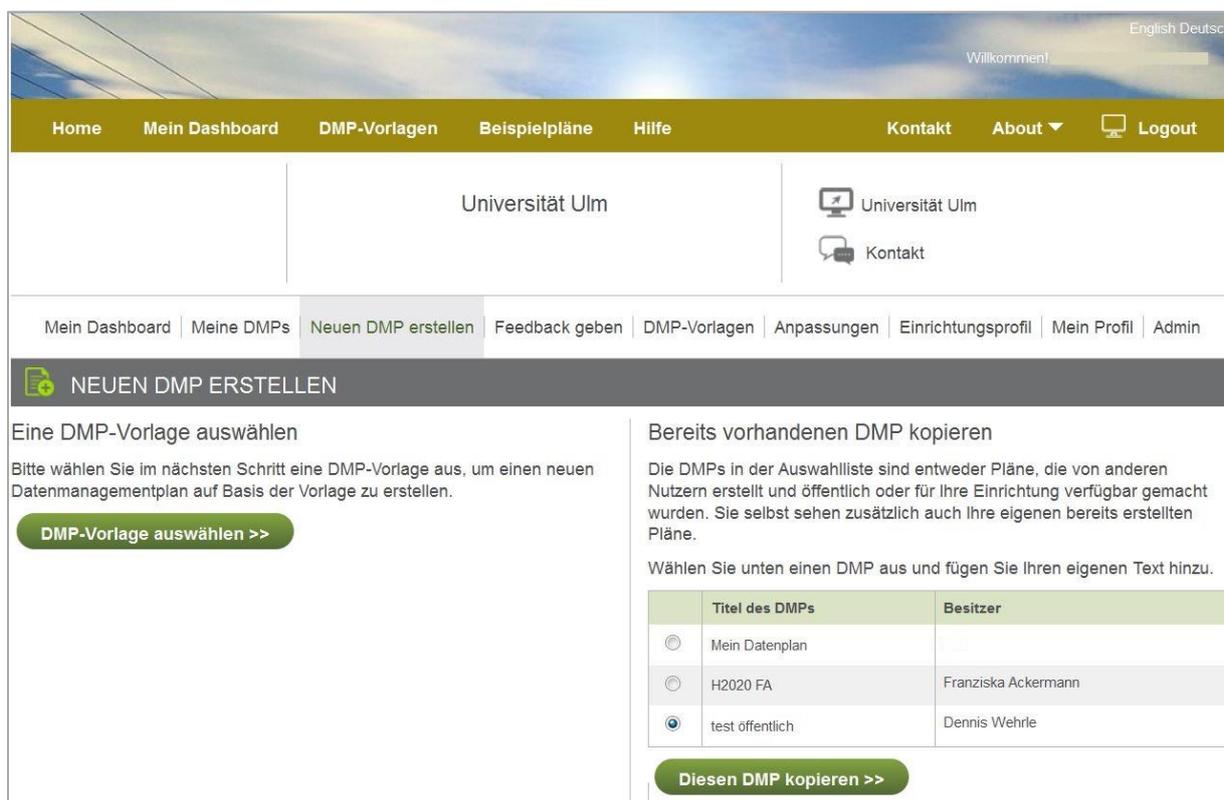


Abb.2: Einen neuen DMP erstellen

8. Fazit

Im Projekt wurde durch die Anpassung des in den USA weit verbreiteten „DMPTool“ eine Lösung für ein kooperatives DMP-Werkzeug geschaffen, das landesweit zum Einsatz kommen und Wissenschaftler bei der Erstellung von Datenmanagementplänen, z.B. für „Horizon 2020“, unterstützen kann. Es soll zunächst an den Standorten Ulm und Freiburg angeboten werden und kann über die Anbindung an Shibboleth / bwIDM darüber hinaus an weiteren Einrichtungen zum Einsatz kommen.

Der Vorteil des webbasierten DMP-Tools besteht unter anderem darin, dass es nur an einer Stelle installiert werden muss und dennoch von einer Vielzahl von Einrichtungen genutzt werden kann. Die Erstellung weiterer Vorlagen und die Anreicherung der Vorlagen durch geeignete Hilfestellungen, Musterantworten sowie einrichtungs- oder fachspezifische Gegebenheiten werden sicherlich eine bedeutende Rolle für die Akzeptanz des DMP-Tools bei den Wissenschaftler darstellen.

Durch den Aufruf der DFG zur Erarbeitung von disziplinspezifischen Data Policies und die zunehmende Wichtigkeit des Themas Forschungsdatenmanagement in Wissenschaft und Politik wird deutlich, dass ein kooperativ nutzbares DMP-Werkzeug ein hilfreicher Baustein in der Forschungsdaten-Infrastruktur des Landes darstellen kann.

Literaturverzeichnis

DFG 2014. Leitfaden für die Antragstellung: Projektanträge.
http://www.dfg.de/formulare/54_01/54_01_de.pdf, Zugriff 25.04.2016

TU Darmstadt, 2015. Leitlinien zum Umgang mit digitalen Forschungsdaten an der TU Darmstadt. http://www.tu-darmstadt.de/media/dezernatvi/relaunch_2015/gute_wiss_praxis/Leitlinien_Forschungsdaten_2015.pdf, Zugriff 10.02.2016

Universität Bielfeld, 2011. Resolution zu Forschungsdatenmanagement. <https://data.uni-bielefeld.de/policy>, Zugriff 10.02.2016

Universität Göttingen, 2014. Forschungsdaten-Leitlinie der Universität Göttingen (einschl. UMG). <http://www.uni-goettingen.de/de/01-juli-2014-forschungsdaten-leitlinie-der-universitaet-goettingen-einschl-umg/488918.html>, Zugriff 10.02.2016

Universität Kiel, 2015. Leitlinie zum Umgang mit Forschungsdaten: Empfehlungen zum Umgang mit Forschungsdaten der Christian-Albrechts-Universität zu Kiel. <http://www.uni-kiel.de/download/pm/2015/2015-408-leitlinie-forschungsdaten.pdf>, Zugriff 10.02.2016

Universität Heidelberg 2014. Research Data Policy: Richtlinien für das Management von Forschungsdaten. <https://www.uni-heidelberg.de/universitaet/profil/researchdata/>, Zugriff 10.02.2016