

Datenmanagementpläne leicht gemacht: TUB-DMP, das Web-Tool der TU Berlin für Datenmanagementpläne

Monika Kuberek und Fabian Fürste, Universitätsbibliothek der Technischen Universität Berlin

Zusammenfassung:

Das Web-Tool TUB-DMP ist ein weiterer Baustein in der Forschungsdaten-Infrastruktur der Technischen Universität Berlin, die vom „Servicezentrum Forschungsdaten und -publikationen“ (SZF) betrieben wird. TUB-DMP ergänzt das Repositorium für Forschungsdaten und Publikationen „Deposit-Once“, das die langfristige Verfügbarkeit, Verifizierung, Zitierbarkeit und Nachnutzbarkeit der Forschungsergebnisse gewährleistet. Forschungsdatenmanagement und Forschungsdaten-Infrastruktur der TU Berlin sind nach den FAIR Data Principles (Findable, Accessible, Interoperable, Re-usable) ausgerichtet. Dieser Maßgabe entsprechend wurde auch TUB-DMP entwickelt, um den Wissenschaftlerinnen und Wissenschaftlern der TU Berlin ein Dokumentationswerkzeug an die Hand zu geben, das ein FAIRes Management ihrer Forschungsdaten ermöglicht. TUB-DMP wurde gemeinsam mit Pilotpartnern aus der TU Berlin entwickelt. Die erste Version wurde im September 2015 in Betrieb genommen, Version 2 mit einer Reihe von Erweiterungen im Dezember 2017.

Summary:

The TUB-DMP Web tool is a further component in the research data infrastructure of the Technische Universität Berlin (TU Berlin), which is operated by the Service Center for Research Data and Publications (SZF). TUB-DMP complements “DepositOnce”, the repository for research data and publications, which guarantees the long-term availability, verification, citability and re-usability of research results. The management of research data and the research data infrastructure at TU Berlin are aligned with the FAIR data principles (Findable, Accessible, Interoperable, Re-usable). TUB-DMP was developed according to these requirements in order to provide scientists with a documentation tool that enables the FAIR data management of their research data. TUB-DMP was developed in cooperation with pilot partners from TU Berlin. The first version went live in September 2015, the extended second version was released in December 2017.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S47-60>

Autorenidentifikation: Kuberek, Monika: ORCID: <https://orcid.org/0000-0002-1672-5271>; Fürste, Fabian: ORCID: <https://orcid.org/0000-0003-2172-609X>

Schlagwörter: Datenmanagementplan, DMP, FAIR Data Principles, Forschungsdatenmanagement, FDM

1. TUB-DMP – ein Web-Tool für Datenmanagementpläne

Das Web-Tool TUB-DMP ist ein weiterer Baustein in der Forschungsdaten-Infrastruktur der Technischen Universität Berlin, die vom Servicezentrum Forschungsdaten und -publikationen (SZF)¹ betrieben wird. TUB-DMP ergänzt das Repositorium für Forschungsdaten und Publikationen DepositOnce,² das die langfristige Verfügbarkeit, Verifizierung, Zitierbarkeit und Nachnutzbarkeit der Forschungsergebnisse (= Forschungsdaten und Publikationen) gewährleistet. Forschungsdatenmanagement (FDM) und Forschungsdaten-Infrastruktur der TU Berlin sind nach den FAIR Data Principles³ (Findable, Accessible, Interoperable, Re-usable) ausgerichtet. Dieser Maßgabe entsprechend wurde auch TUB-DMP entwickelt, um den Forschenden an der TU Berlin ein Dokumentationswerkzeug an die Hand zu geben, das ein FAIRes Management ihrer Forschungsdaten ermöglicht.

1.1. Was ist ein Datenmanagementplan (DMP)?

Ein DMP trägt dazu bei, dass Forschungsdaten langfristig ihren Wert behalten. Er dokumentiert den Kontext, in dem die Forschungsdaten entstanden sind, sowie den Umgang mit den Forschungsdaten während ihrer Entstehung und nach Beendigung ihrer Erhebung. Der DMP enthält Angaben zu Projektleitung, Projekttitel, Projektbeschreibung, Förderorganisation, Art und Umfang der entstandenen Forschungsdaten, Speicherung, Aufbewahrungsdauer usw. Diese Informationen dienen insbesondere dazu, Forschungsdaten in späteren Jahren interpretieren, nachvollziehen und somit nachnutzen zu können.

Die Verfügbarkeit gut organisierter und gepflegter Daten gehört heute zur guten wissenschaftlichen Praxis und gewährleistet die Transparenz der Forschungsergebnisse, eine grundlegende Anforderung im Forschungsbereich.

1.2. Ausgangslage

Im letzten Jahrzehnt ist weltweit die Erkenntnis gewachsen, dass Forschungsdaten eine unschätzbare Ressource und einen Wert an sich darstellen, den es zu erhalten, disziplinübergreifend bereitzustellen und nachzunutzen gilt. Das in der TU Berlin in den Jahren 2012–2014 aufgebaute und seitdem etablierte FDM umfasst den gesamten Lebenszyklus der Forschungsdaten und steht mit seiner konzeptionellen Zielstellung – Speicherung, Verifizierung, Langzeitarchivierung und Nachnutzung von Forschungsergebnissen – auch dafür ein, dass die Forschungsdaten langfristig ihren Wert als Forschungsressource behalten.⁴ So sollen in DepositOnce, dem Repositorium der TU Berlin, zusammen mit den Forschungsdaten auch Informationen gespeichert werden, die zur Verifizierung und Reproduktion eines Forschungsergebnisses notwendig sind.⁵

1 Das SZF ist eine gemeinsame Einrichtung der Universitätsbibliothek (UB), des IT-Service-Center tubIT und der Forschungsabteilung (Abt. V); die Leitung des SZF ist in der Hauptabteilung 3 „Elektronische Dienste“ (HA 3) der UB angesiedelt, zuletzt geprüft am 15.02.2018, http://www.szf.tu-berlin.de/menue/ueber_das_szf.

2 Repositorium DepositOnce, zuletzt geprüft am 15.02.2018, <https://depositonce.tu-berlin.de>.

3 Mark D. Wilkinson et al., „The FAIR Guiding Principles for scientific data management and stewardship,“ *Sci. Data* 3:160018 (2016), <https://doi.org/10.1038/sdata.2016.18>.

4 Monika Kuberek, *Organisatorisch-technisches Konzept für eine Forschungsdaten-Infrastruktur in der TU Berlin*. Technische Universität Berlin, 15. März 2012, 4–5, <http://dx.doi.org/10.14279/depositonce-6604>.

5 Ebd., 14–15.

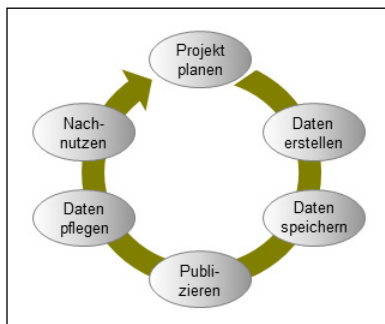


Abb. 1: Forschungsdaten-Lebenszyklus (Grafik angelehnt an: <https://www.ukdataservice.ac.uk/manage-data/lifecycle>)

Darüber hinaus sind für die Nachnutzung der Forschungsdaten weitere Informationen unerlässlich, wie zum Beispiel: Wie und in welchem Rahmen / Projekt sind die Forschungsdaten entstanden? Welche Methoden wurden bei der Erstellung angewendet? Welches ist der Erhebungszeitraum? Wie und wo werden die Daten gespeichert? Welche Kosten entstehen langfristig? Der Bereitstellung dieser Informationen dient der DMP.

Vor dem Hintergrund der Nachnutzung von Forschungsdaten gehört die Erstellung von DMPs in den anglo-amerikanischen Ländern seit etlichen Jahren zum Standard bei der Projektantragstellung; sie werden in aller Regel von den dortigen Förderorganisationen eingefordert. Ebenso wie beim FDM insgesamt kommt Großbritannien auch bei der Entwicklung und dem Einsatz von DMPs die Vorreiterrolle in Europa und eine führende Rolle weltweit zu. Der erste standardisierte DMP⁶ in Großbritannien wurde 2009 vom Digital Curation Centre (DCC)⁷ entwickelt und seitdem kontinuierlich weiterentwickelt. Die DCC-Checklist, die inzwischen in der 4. Version⁸ vorliegt, ist die in der Wissenschaftslandschaft international am weitesten verbreitete Vorlage zur Erstellung eines DMPs; sie war auch das Vorbild für die im DFG-Projekt WissGrid⁹ entwickelte „Checkliste zum Forschungsdaten-Management“.¹⁰

- 6 Martin Donnelly und Sarah Jones, *Data Management Plan Content Checklist: Draft Template for Consultation* (Edinburgh: Digital Curation Centre, 17.06.2009), zuletzt geprüft am 15.02.2018, http://www.dcc.ac.uk/webfm_send/1280. Siehe auch: *Checklist for a Data Management Plan*, zuletzt geprüft am 15.02.2018, <http://www.dcc.ac.uk/resources/data-management-plans/checklist>.
- 7 DCC wurde 2004 auf Empfehlung von JISC (<https://www.jisc.ac.uk/>), dem damaligen „Joint Information Systems Committee“ als nationales Zentrum zur Sicherung des dauerhaften Zugangs und der Langzeitarchivierung digitaler Ressourcen gegründet, zuletzt geprüft am 15.02.2018, <http://www.dcc.ac.uk/about-us/history-dcc/history-dcc>.
- 8 DCC, *Checklist for a Data Management Plan*, v4.0 (Edinburgh: Digital Curation Centre, 2013), zuletzt geprüft am 15.02.2018, http://www.dcc.ac.uk/webfm_send/1279.
- 9 Das DFG-Projekt „Grid für die Wissenschaft – WissGrid“ unter Leitung der Georg-August Universität Göttingen wurde am 30.04.2012 nach dreijähriger Laufzeit abgeschlossen, zuletzt geprüft am 15.02.2018, <https://www.sub.uni-goettingen.de/projekte-forschung/projekt/details/projekt/wissgrid/>.
- 10 Jens Ludwig und Harry Enke, Hrsg., *Leitfaden zum Forschungsdaten-Management – Handreichungen aus dem WissGrid-Projekt* (Glückstadt: vwh, 2013), 83ff, zuletzt geprüft am 15.02.2018, http://www.forschungsdaten.org/images/b/b0/Leitfaden_Data-Management-WissGrid.pdf.

Die Nachnutzung von Forschungsdaten ist in den vergangenen Jahren auch in Deutschland immer stärker in den Fokus der Drittmittelgeber gerückt: Die Deutsche Forschungsgemeinschaft (DFG) verlangt seit einigen Jahren bei der Antragstellung entsprechende Konzepte zur Projektdokumentation,¹¹ hat diese auf ihrer Webseite „Umgang mit Forschungsdaten“¹² in den übergeordneten Leitlinien für Antragstellende dargelegt und die Forderungen 2015 in einer Leitlinie verankert.¹³ Das Bundesministerium für Bildung und Forschung (BMBF) macht die Erstellung von DMPs nicht generell obligatorisch für die Antragstellung, fordert allerdings in einer Reihe von Programmen – insbesondere, wenn es sich um transnationale Programme handelt – die Erstellung von DMPs.¹⁴

Das EU-Rahmenprogramm für Forschung und Innovation für die Jahre 2014–2020 „Horizon 2020“ ging einen Schritt weiter als die deutschen Förderorganisationen und machte bereits in der ersten Förderperiode ab 2014 die Erstellung eines DMPs für alle am Open Research Data (ORD) Pilot beteiligten Forschungsprojekte obligatorisch: „Das Pilotprojekt für offene Forschungsdaten betrifft sowohl die Grunddaten, die zur Reproduktion der in wissenschaftlichen Veröffentlichungen vorgestellten Ergebnisse nötig sind, als auch andere Daten, die von den Forschern in ihren Datenmanagementplänen (DMP) angegeben werden. Dies ist ein neues Element in ‚Horizont 2020‘. Ein DMP legt einen Datenverwaltungszyklus für alle Daten fest, die im Zuge des Forschungsprojekts erfasst, verarbeitet oder erzeugt werden. Alle Forschungsprojekte, die sich an dem Pilotprojekt beteiligen, müssen einen DMP haben, die anderen ‚Horizont-2020‘-Projekte sind aufgefordert, freiwillig einen solchen Plan aufzustellen, wenn dies für ihre Forschung relevant ist.“¹⁵ In der Fortschreibung des Rahmenprogramms 2017 wurde der ORD Pilot von den ursprünglich sieben thematischen Förderbereichen auf alle Themengebiete von Horizon 2020 ausgeweitet. In den aktuellen „Guidelines on FAIR Data Management in Horizon 2020“ vom Juli 2017 ist der DMP als „key element of good data management“¹⁶ definiert und erstmals die Vorlage für einen DMP, das „Horizon 2020 FAIR Data Management Plan (DMP) template“,¹⁷ als Anhang beigelegt. Während es für die Teilnehmer am ORD Pilot hinsichtlich der Veröffentlichung ihrer Forschungsdaten verschiedene opt-out-Möglichkeiten gibt,¹⁸ heißt es zur Erstellung eines DMP: „A DMP is required for all projects participating in the extended ORD pilot, unless they opt out of the ORD pilot. However, projects that opt out are still encouraged to submit

11 DFG, *Leitfaden für die Antragstellung – Projektanträge* (Bonn: DFG, 2018), 5–6, zuletzt geprüft am 15.02.2018, http://www.dfg.de/formulare/54_01/54_01_de.pdf.

12 DFG, *Umgang mit Forschungsdaten*. DFG-Leitlinien zum Umgang mit Forschungsdaten, Aktualisierungsdatum: 29.10.2015 (Bonn: DFG, 2015), zuletzt geprüft am 15.02.2018, http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/.

13 DFG, *Leitlinien zum Umgang mit Forschungsdaten*, verabschiedet durch den Senat der DFG am 30.09.2015, (Bonn: DFG, 2015), zuletzt geprüft am 15.02.2018, www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf.

14 Siehe beispielsweise: BMBF, „Bekanntmachung: Richtlinie zur Förderung transnationaler Verbundvorhaben auf dem Gebiet der Systemmedizin innerhalb des ERA-Netzes ERACoSysMed,“ *Bundesanzeiger* (03.02.2017), Punkt 2, Datenmanagement, zuletzt geprüft am 15.02.2018, <https://www.bmbf.de/foerderungen/bekanntmachung-1309.html>.

15 Siehe Pressemitteilung zum Start des Pilotprojekts: Europäische Kommission, „Kommission startet Pilotprojekt zur Öffnung der mit öffentlicher Förderung erzeugten Forschungsdaten,“ *Pressemitteilung* (16. Dezember 2013), zuletzt geprüft am 15.02.2018, http://europa.eu/rapid/press-release_IP-13-1257_de.htm.

16 European Commission, *H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020*, Version 3.0, 26. July 2016, (Brussels: European Commission, 2016), 4, zuletzt geprüft am 15.02.2018, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.

17 Ebd., 6–9.

18 Ebd., 3.

a DMP on a voluntary basis.”¹⁹ Mit ihrem Förderprogramm Horizon 2020 hat die EU-Kommission seit 2014 entscheidende Anreize gesetzt, dass mit öffentlicher Förderung erzeugte Forschungsdaten frei zugänglich gemacht werden. Gleichzeitig hat sie mit der engen Kopplung von Forschungsprojektförderung und DMP und der stetigen Argumentation, dass zusätzliche Informationen zu den Forschungsdaten für deren Nachnutzung zwingend notwendig sind, das Verständnis gefördert, dass zu jedem Forschungsprojekt ein DMP gehört.

Horizon 2020 gab auch an der TU Berlin wesentliche Impulse für die Erstellung von DMPs. In Zusammenarbeit mit Pilotpartnern wurde im Rahmen des SZF ein Web-Tool zur Erstellung von DMPs, TUB-DMP,²⁰ entwickelt und Version 0 im August 2014 für einen erweiterten Kreis zum Probetrieb freigegeben.²¹ Im September 2015 wurde Version 1 für den Routinebetrieb zur Nutzung freigeschaltet und ist seitdem an der TU Berlin im regulären Einsatz.²² Seit Dezember 2017 läuft Version 2 von TUB-DMP im Routinebetrieb, die eine Reihe von Erweiterungen enthält, wie Login mittels TUB-Account, Anlegen beliebig vieler DMPs für Teilprojekte von Verbundvorhaben, Erstellung von DMPs für Vorhaben ohne TUB-Projektnummer. TUB-DMP kann direkt oder auf den Seiten des SZF²³ in der Produktionsversion und in einer Demoversion aufgerufen werden.

1.3. Mehrwert für die Forschenden – das Konzept von TUB-DMP

Neben Beratung und Information besteht die grundlegende Zielsetzung des SZF darin, Mehrwertdienste für die Forschenden zu schaffen, um ihnen den Umgang mit Forschungsergebnissen zu vereinfachen, Arbeitsweisen zu effektivieren und die tägliche Arbeit zu erleichtern. Zur Entwicklung bedarfsgerechter, leicht bedienbarer Werkzeuge arbeitet das SZF mit Pilotpartnern aus dem Kreis der Forschenden der TU Berlin zusammen. Diesem Leitgedanken ist das SZF auch bei der Entwicklung des Web-Tools TUB-DMP gefolgt: Mit TUB-DMP wird die Erstellung eines DMPs leicht gemacht.

TUB-DMP wurde als Mehrwertdienst konzipiert und entwickelt:

- Bereits vorhandene Daten aus dem Projektantrag werden automatisch übernommen.
- Informationen zu DepositOnce, dem Repositorium der TU Berlin, sind vorab bereits eingetragen und müssen nicht mehr ermittelt werden (DepositOnce-Standardwerte).
- Der Step-by-Step-Workflow leitet die Forschenden auf einfache Weise durch einen Fragenkatalog. Er fragt alle wesentlichen Bestandteile eines DMPs ab und sorgt dafür, dass nichts Wichtiges vergessen wird.

19 Ebd., 4.

20 „TUB-DMP,“ zuletzt geprüft am 15.02.2018, <https://dmp.tu-berlin.de>.

21 Das Tool wurde erstmals auf dem Bibliothekartag 2014 in Bremen vorgestellt. Siehe Fabian Fürste, „TUB-DMP – Der Datenmanagementplan als Bindeglied zwischen Forschungsinformationssystem und Repositorium,“ (Vortrag auf dem 103. Deutschen Bibliothekartag in Bremen, 2014), Vortragsfolien unter: <https://nbn-resolving.org/urn:nbn:de:0290-opus-17540>. Siehe auch: Fabian Fürste, „TUB-DMP – Ein Werkzeug zur Erstellung von Datenmanagementplänen an der TU Berlin,“ in DINI-Workshop „Datenmanagementpläne und ihre Bedeutung im Forschungsdatenmanagement“ (Berlin, 2015), zuletzt geprüft am 11.04.2018, <http://www.forschungsdaten.org/images/d/dc/06--fuerst-werkzeug-erstellung-datenmanagementplaene.pdf>.

22 Monika Kuberek, „Der Umgang mit Forschungsdaten: Für ‚Horizon 2020‘-Anträge müssen Datenmanagementpläne entwickelt werden; Die Universitätsbibliothek hat ein neues Tool entwickelt, um Forschende zu unterstützen,“ TU intern Nr. 7-9 (2016), 4, zuletzt geprüft am 11.04.2018, <http://archiv.pressestelle.tu-berlin.de/tui/16juli/#/4>.

23 „Datenmanagementplan TUB-DMP,“ zuletzt geprüft am 11.04.2018, http://www.szf.tu-berlin.de/menue/dienste_tools/datenmanagementplan_tub_dmp.

- Im Eingabeformular werden die Daten strukturiert abgelegt und können – übersichtlich gegliedert – als Ausdruck an die Förderorganisation weitergegeben werden.
- Verschiedene Templates, so unter anderem zur Erstellung eines Horizon 2020-konformen DMPs, stehen zur Auswahl.
- In TUB-DMP können DMPs für Einzelprojekte wie auch für Verbundprojekte mit vielen Teilprojekten angelegt werden.
- Die Forschenden können sich mit ihrem TUB-Account in TUB-DMP anmelden.
- TUB-DMP verfügt über ein Helpdesk (Ticketsystem), an das sich die Forschenden bei Problemen oder Fragen wenden können. Für die Beantwortung erster Fragen zu TUB-DMP gibt es ausführliche FAQs²⁴ auf den Seiten des SZF. Das SZF-Team bietet darüber hinaus bei Bedarf Beratungen und Workshops an.

Durch die Nachnutzung der Projektantragsdaten und durch die DepositOnce-Standardwerte sind beim DCC-Template fast 30%, beim weniger umfangreichen Horizon2020-Template sogar rund 40% des DMP bereits ausgefüllt, wenn der Plan zum ersten Mal aufgerufen wird. Mit diesem konzeptionellen Ansatz des Mehrwertdienstes unterscheidet sich TUB-DMP grundlegend von dem weit verbreiteten Web-Tool DMPonline,²⁵ in dem alle Informationen manuell per Fließtext eingegeben werden müssen. Der seit November 2015 im Rahmen eines DFG-Projekts entwickelte Research Data Management Organiser (RDMO) bietet in seinem jetzigen Entwicklungsstand weder die Mehrwertdienste noch den Funktionsumfang von TUB-DMP.²⁶

TUB-DMP unterscheidet drei Kategorien von Daten und Eingabefeldern:

a) Daten, die bereits in der Forschungsprojektdatenbank der TU Berlin gespeichert sind

Um Doppelerfassung zu vermeiden und zur Arbeitserleichterung werden bereits vorhandene Projektinformationen aus der Antragstellung (wie Projekttitel, Abstract, Projektleitung, Projektlaufzeiten,...) in TUB-DMP nachgenutzt. Die Informationen werden bei der Initiierung eines Plans übertragen und liegen bereits vor, wenn eine Wissenschaftlerin oder ein Wissenschaftler ihn erstmals aufruft.

b) DepositOnce-Standardwerte

Als besonderen Mehrwert für die Forschenden, die ihre Forschungsdaten in DepositOnce speichern, gibt es die DepositOnce-Standardwerte. Das sind Kurztexte mit Informationen zu Deposit-Once (wie Infos zu den Betreibern des Repositoriums und zur Infrastruktur, Sicherheitskonzept, Speicherzyklen, usw.), die das SZF-Team vorab erstellt und in der TUB-DMP-Datenbank gespeichert hat. Auf diese Weise sind die entsprechenden Fragen in TUB-DMP bereits mit Antworten versehen, wenn

24 Nutzung von TUB-DMP, zuletzt geprüft am 11.04.2018, http://www.szf.tu-berlin.de/menue/faq/nutzung_von_tub_dmp.

25 DMPonline ist eine gemeinsame Entwicklung des Digital Curation Centre (DCC) in Großbritannien und des University of California Curation Center (UC3). Der Quellcode ist open source und auf GitHub eingestellt, zuletzt geprüft am 11.04.2018, <https://dmponline.dcc.ac.uk>.

26 Im November 2017 ist die zweite Projektförderphase von RDMO angelaufen, in deren Verlauf „... u.a. Rollenkonzepte und Anbindungen an die institutionelle Infrastruktur wie Repositorien, Ticket-Systeme und Infrastruktur für Authentifizierung und Autorisierung“ entwickelt werden sollen, wie es auf der RDMO-Website heißt, zuletzt geprüft am 11.04.2018, <https://rdmorganiser.github.io>.

Forschende ihren Plan zum ersten Mal aufrufen. Sie ersparen sich damit das Erfragen dieser Informationen bei den Betreibern von DepositOnce – eine nicht unwesentliche Arbeitserleichterung.

c) Sonstige Inhalte, die sich auf das Projekt oder die Forschungsergebnisse beziehen

Darüber hinaus ermöglicht TUB-DMP die strukturierte Eingabe weiterer Informationen in definierten Freitextfeldern.

Alle Inhalte in TUB-DMP können von den Forschenden geändert werden, einschließlich der Werte aus der Forschungsprojektdatenbank und der DepositOnce-Standardwerte, die dem spezifischen Projekt entsprechend angepasst werden können.

2. Struktur und Inhalt von TUB-DMP

In TUB-DMP können unterschiedliche Templates implementiert werden, um DMPs für spezielle Förderprogramme oder verschiedene Disziplinen zu erstellen. Als Basis-Template ist ein allgemeines Template hinterlegt, dessen Inhalt der DCC-Checklist²⁷ entspricht. Anhand des allgemeinen Templates sollen hier Struktur und Inhalt von TUB-DMP erläutert werden.

Die DCC-Checklist untergliedert sich in Sektionen: Auf die Sektion „Administrative Data“ mit allgemeinen Angaben zum Projekt folgen die Sektionen 1 bis 8, die sich auf die Forschungsdaten selbst beziehen. Entsprechend der DCC-Sektionen sind die Felder im TUB-DMP-Eingabeformular untergliedert und benannt. Das strukturierte Eingabeformular mit den vorformulierten Fragen bildet eine Art „Gerüst“ und soll den Forschenden eine Anleitung und Hilfestellung geben, welche Informationen ein DMP enthalten soll. Nicht alle Felder in TUB-DMP müssen ausgefüllt werden. Leere Felder werden beim Export ignoriert, sodass der Ausdruck – beispielsweise zur Vorlage bei einer Förderorganisation – nur die Informationen aus den ausgefüllten Feldern enthält.

Um Verbundprojekte abzubilden, verfügt TUB-DMP über eine hierarchische Struktur. Verbundvorhaben (wie DFG-Sonderforschungsbereiche, DFG-Forschergruppen, DFG-Graduiertenkollegs, BMBF- und EU-Verbundprojekte) bestehen in der Regel aus mehreren Teilprojekten, die jeweils eigene Inhalte und Zielsetzungen haben und in denen jeweils unterschiedliche Forschungsdaten generiert werden. Jedes dieser Teilprojekte benötigt einen eigenen DMP. In TUB-DMP ist es möglich, DMPs für beliebig viele Teilprojekte eines Verbundprojekts anzulegen. Dazu wurde eine hierarchische Struktur realisiert, die für Einzelprojekte wie auch für Verbundprojekte genutzt wird: Auf der obersten Ebene gibt es den „Projekt-DMP“,²⁸ darunter pro Teilprojekt einen eigenen „DMP“. Zu einem „Projekt-DMP“ können beliebig viele „DMPs“ angelegt werden. Das Verhältnis „Projekt-DMP“:„DMP“ ist bei einem Einzelprojekt 1:1, bei einem Verbundvorhaben 1:n.

²⁷ DCC-Checklist.

²⁸ Die Bezeichnungen „Projekt-DMP“ und „DMP“ sind eventuell etwas verwirrend und seien hier kurz erläutert: Bei der Implementierung der hierarchischen Struktur mussten unterschiedliche Bezeichnungen für die beiden Hierarchiestufen gewählt werden; die Entscheidung fiel auf „Projekt-DMP“ für die oberste und „DMP“ für die darunterliegende Hierarchiestufe. Diese technischen Begriffe und ihre Unterscheidung sind nur im Zusammenhang mit der hierarchischen Struktur von Bedeutung und hier in Anführungszeichen gesetzt. Ansonsten ist in diesem Artikel mit der Abkürzung DMP immer die Gesamtheit des Plans gemeint, der sich aus „Projekt-DMP“ und „DMP“ zusammensetzt.

Der „Projekt-DMP“ enthält die übergeordneten Projektinformationen (wie Titel, Abstract, Projektleitung usw.) entsprechend der DCC-Sektion „Administrative Data“. Der „DMP“ enthält die Informationen zu den Forschungsdaten, entsprechend der DCC-Sektionen 1–8. Die Inhalte des „Projekt-DMP“ werden bei der Initiierung des Plans aus der Forschungsprojektdatenbank übernommen. Im „DMP“ finden sich dann die vorab generierten DepositOnce-Standardwerte und weitere Textfelder, die von den Wissenschaftlerinnen und Wissenschaftlern manuell ausgefüllt werden.

3. Der Workflow in TUB-DMP

Das Schema in Abbildung 2 zeigt den Workflow in TUB-DMP – inklusive der Akteure (Wissenschaftlerinnen und Wissenschaftler sowie SZF-Team) auf der linken Seite und den Datenquellen / Schnittstellen zu TUBIS²⁹ (Authentifizierung) sowie zur Forschungsprojektdatenbank (Übernahme der Projektdaten) auf der rechten Seite.

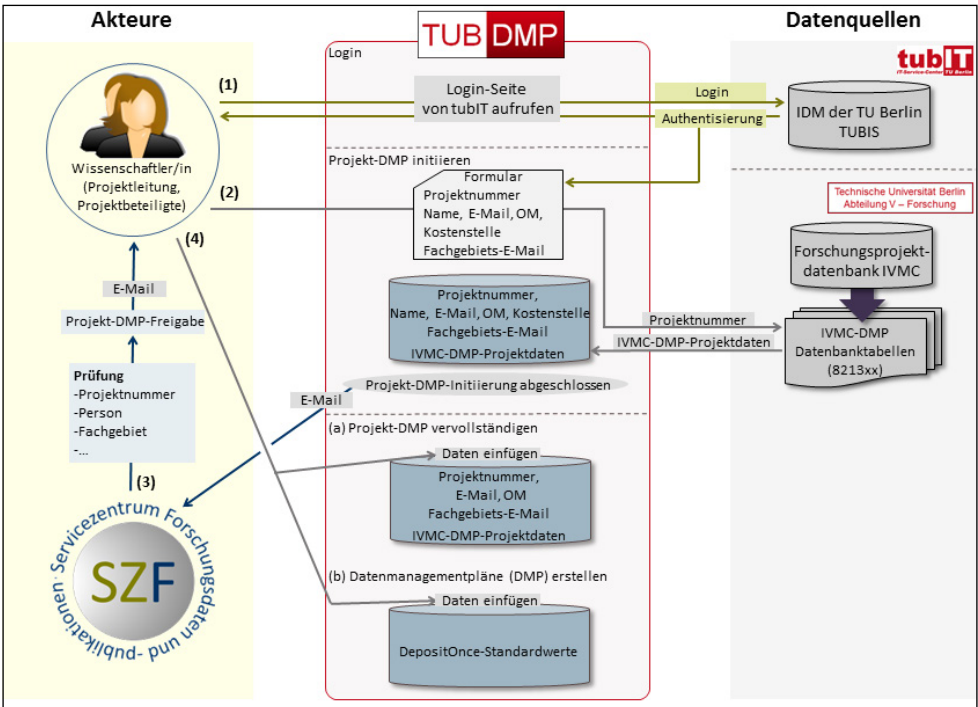


Abb 2: Schematische Darstellung des Workflows in TUB-DMP

Im Folgenden werden die Schritte im Einzelnen beschrieben; das Verfahren ist für Einzelprojekte und für Verbundprojekte mit vielen Teilprojekten identisch:

²⁹ TUBIS (TU Berlin Integrations Service) ist das von tubIT betriebene, rollenbasierte Identitätsmanagement der TU Berlin.

a) Authentifizierung (Wissenschaftler/innen)

Beim Anklicken des Login-Buttons auf der TUB-DMP-Homepage werden die Wissenschaftlerinnen und Wissenschaftler über zur Loginseite von tubIT weitergeleitet und authentifizieren sich dort mit ihrem TUB-Account.

b) Plan initiieren (Wissenschaftler/innen)

Nach dem Login erhalten sie in TUB-DMP ein Formular mit einigen Authentifizierungsmerkmalen aus TUBIS (wie Name, Fachgebiet), das sie mit der TUB-Projektnummer (so vorhanden) und der institutionellen E-Mail-Adresse des Fachgebiets vervollständigen. Sobald das Formular abgeschickt wird, wird ein „Projekt-DMP“ generiert.

- Ist das Feld mit der Projektnummer ausgefüllt und war die Anfrage in den IVMC-DMP-Datenbanktabellen positiv, werden die Projektdaten in die TUB-DMP-Datenbank eingefügt und stehen dann den Wissenschaftlerinnen und Wissenschaftlern in den entsprechenden Feldern des „Projekt-DMP“ zur Bearbeitung zur Verfügung.
- Ist das Feld mit der Projektnummer ausgefüllt und war die Anfrage in den IVMC-DMP-Datenbanktabellen negativ, wird ein leerer „Projekt-DMP“ mit der angegebenen Projektnummer generiert.
- Ist das Feld mit der Projektnummer leer, wird eine zufällig generierte Projektnummer vom System erzeugt und in der TUB-DMP-Datenbank gespeichert.

Beim Abschicken des Formulars wird zudem das SZF-Team in der UB per E-Mail darüber informiert, dass ein neuer „Projekt-DMP“ angelegt wurde.

c) Formalprüfung (UB)

Bevor der Plan von den Wissenschaftlerinnen und Wissenschaftlern bearbeitet werden kann, erfolgt eine formale Prüfung durch das SZF-Team, ob seine Initiierung plausibel war (anhand der Einträge im „Projekt-DMP“, Zugehörigkeit der eintragenden Person zum Fachgebiet usw.). Ist die Plausibilität gegeben, wird der Plan vom SZF-Team freigegeben und eine E-Mail generiert, mit der die Person, die den Plan initiiert hat, von der Freigabe in Kenntnis gesetzt wird.

d) Vervollständigung der Daten (Wissenschaftler/innen)

Nach Freigabe des „Projekt-DMP“ können die Wissenschaftlerinnen und Wissenschaftler diesen, falls notwendig, vervollständigen und einen „DMP“ mit den projektspezifischen Informationen anlegen.

Alle Änderungen werden in der TUB-DMP-Datenbank gespeichert. Zur weiteren Verwendung können die Wissenschaftlerinnen und Wissenschaftler den Datensatz als PDF exportieren.

Falls Änderungen vorgenommen werden sollen, kann ein DMP über das Login jederzeit wieder aufgerufen und bearbeitet werden.

4. Technische Umsetzung

4.1. Das Web-Tool TUB-DMP

TUB-DMP ist eine Eigenentwicklung der UB. Sie übernimmt im Rahmen des SZF auch die Pflege des Web-Tools.

a) Software

Das Web-Tool setzt auf dem modernen PHP-Framework Laravel auf, welches wiederum auf dem Symfony-Framework basiert. TUB-DMP wird über den Apache-Webserver ausgeliefert. Die DMPs werden in einer relationalen Datenbank gespeichert und verwaltet (TUB-DMP-Datenbank); als Datenbanksoftware wird PostgreSQL verwendet. Die Anwendung nutzt weitverbreitete Tech Stacks und stellt keine besonderen Anforderungen an Hard- und Software. Der Quellcode von TUB-DMP wird auf GitHub als Open Source Software veröffentlicht.³⁰

b) Integration in die IT-Infrastruktur der TU Berlin

Die schematische Darstellung in Abbildung 3 zeigt die Integration von TUB-DMP in die IT-Infrastruktur der TU Berlin, die vom Rechenzentrum der TU Berlin, dem IT-Service-Center tubIT, betrieben wird. Auf diese Weise sind zum einen eine hohe Ausfallsicherheit sowie ein hoher Schutz der Anwendung und der Daten gewährleistet, zum anderen werden durch die Nachnutzung der vorhandenen Infrastruktur Synergien genutzt, die den Betriebsaufwand für TUB-DMP verringern.

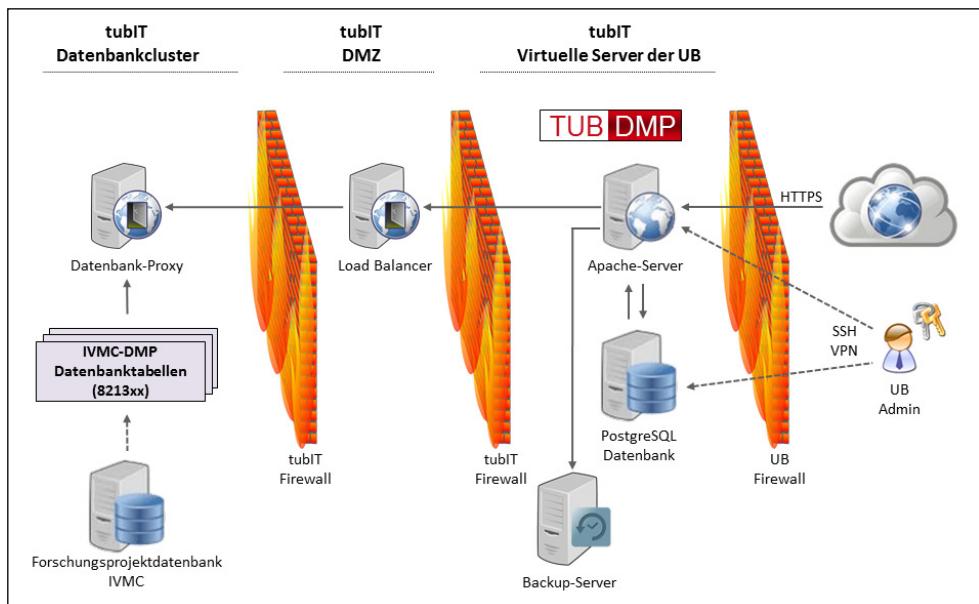


Abb 3: Integration von TUB-DMP in die IT-Infrastruktur der TU Berlin

30 In Vorbereitung. GitHub-Account der UB der TU Berlin, zuletzt geprüft am 11.04.2018, <https://github.com/tuub>.

Die Datenbank und das Dateisystem von TUB-DMP laufen auf virtuellen Servern der UB, die Teil der von tubIT betriebenen virtualisierten Serverumgebung der TU Berlin sind. Die in TUB-DMP eingegebenen Daten werden im Backup-System der TU Berlin, das ebenfalls von tubIT betrieben wird, gesichert. Der Zugriff von außen auf TUB-DMP ist durch verschlüsselte Netzwerkverbindungen geschützt: Nur die TUB-DMP-Administratorinnen und -Administratoren der UB haben im Campusnetz der TU Berlin – mittels SSH (Secure Shell) in Verbindung mit VPN (Virtual Private Network) – Zugriff auf die Datenbank und das Dateisystem. Für den Web-Zugang der Benutzerinnen und Benutzer auf TUB-DMP ist das sichere Kommunikationsprotokoll HTTPS (Hypertext Transfer Protocol Secure) implementiert.

Der Abruf der Projektdaten aus der Forschungsprojektdatenbank, um sie in TUB-DMP zu übernehmen, erfolgt über eine spezielle Kommunikationsschnittstelle, einen sogenannten Proxy. Dieser, wie auch die Forschungsprojektdatenbank, die sensible Daten enthält, liegen in einem besonders sicheren Bereich von tubIT – geschützt durch eine DMZ (Demilitarized Zone). Um zu verhindern, dass die Forschungsdatenbank durch eine zu große Anzahl von Anfragen beeinträchtigt wird, ist ein sogenannter Load-Balancer zwischengeschaltet. Aus Datenschutzgründen erfolgt von TUB-DMP kein direkter Zugriff auf die Forschungsprojektdatenbank, sondern lediglich auf einen eigens für TUB-DMP definierten Auszug, der in den IVMC-DMP-Datenbanktabellen abgelegt ist. Das Verfahren zur Übernahme der Projektdaten und weitere Erläuterungen zur Skizze sind im nachfolgenden Kapitel beschrieben.

4.2. Automatisierte Übernahme von Projektdaten nach TUB-DMP

Die Forschungsprojektdatenbank IVMC (Integriertes Verwaltungs- und Management Controlling) ist die zentrale Datenbank der TU Berlin, in der die Informationen zu ihren Forschungsprojekten gespeichert werden. Für den DMP werden nicht die Daten aus der ursprünglichen Projektantragstellung verwendet, sondern die Daten, die nach der Projektbewilligung von der Forschungsabteilung (Abt. V) aktualisiert, gegebenenfalls korrigiert oder auch ergänzt wurden. Die Übernahme der Projektdaten aus IVMC nach TUB-DMP wurde in Zusammenarbeit der drei Kooperationspartner im SZF – UB, tubIT und Abt. V – konzipiert und umgesetzt. Eindeutiger Identifikator ist die TUB-Projektnummer, die in der Forschungsprojektdatenbank hinterlegt ist.

Aus Datenschutzgründen darf es keinen direkten Zugriff von TUB-DMP auf die IVMC-Datenbank geben. Daher wurden von UB und Abt. V die für die Nachnutzung relevanten Felder definiert – eine kleine, eng gefasste und fest definierte Untermenge der Forschungsprojektdatenbank mit allgemeinen Projektinformationen. Zur Datenübernahme werden diese Daten in separaten Datenbanktabellen (IVMC-DMP-Datenbanktabellen) auf dem IVMC-Datenbankserver abgelegt und in einem spezifischen Nummernkreis (8213xx) bereitgestellt. Für die Transaktion ist die UB mit einem Service-Account gegenüber dem IVMC-Datenbankserver authentifiziert. Während des Anfrage-Vorgangs ist die Verbindung zwischen dem TUB-DMP-Server und der IVMC-DMP-Datenbank durch eine Transportverschlüsselung gesichert.

4.3. Funktionen von TUB-DMP

a) Zugriffsschutz und Authentifizierung über Shibboleth

Das Konzept von TUB-DMP sieht vor, dass die Forschenden ihre DMPs in einem zugriffsgeschützten Bereich erstellen und bearbeiten. In Zusammenarbeit mit tubIT hat die UB in TUB-DMP Shibboleth implementiert, so dass sich die Wissenschaftlerinnen und Wissenschaftler mittels ihres TUB-Accounts authentifizieren können.³¹

b) Nutzung unterschiedlicher Templates

Es werden Templates für verschiedene Zwecke zur Verfügung gestellt, die per Mausklick aus einer Liste ausgewählt werden können. Zurzeit ist neben dem allgemeinen Template ein Template zur Erstellung Horizon 2020-konformer DMPs implementiert. Es basiert auf dem „Horizon 2020 FAIR Data Management Plan (DMP) template.“³² Weitere Templates sind in Vorbereitung.

c) Speicherfunktion und neue Versionen

Jeder DMP soll bei Bedarf, dem Projektfortgang entsprechend, geändert oder aktualisiert werden. Damit wird den Anforderungen von Horizon 2020 entsprochen, wo vorgesehen ist, dass der DMP im Laufe des Projektfortschritts weitergeschrieben und aktualisiert wird.³³

Während der Eingabe in TUB-DMP kann die Arbeit jederzeit unterbrochen und der Plan zwischengespeichert werden. Die Einträge bleiben erhalten und der Plan kann beim erneuten Aufruf weiterbearbeitet werden. Beim Abspeichern wird jeweils das Datum der letzten Änderung automatisch erstellt und gespeichert. Ist die Erstellung eines Plans abgeschlossen, kann er mit Klick auf einen spezifischen Button als Version abgespeichert werden; danach ist die Bearbeitung nicht mehr möglich. Soll der DMP geändert werden, kann beim nächsten Aufruf eine neue Version angelegt werden. Diese enthält die Daten der Vorgängerversion, die geändert und bearbeitet werden können.

Bei der Ersterstellung eines DMP wird das Erstelldatum automatisch generiert und gespeichert; es bleibt als konstantes Datum immer erhalten. Im DMP werden das Datum der Ersterstellung und das Datum der letzten Änderung angezeigt. Für die bearbeitende Person sind in TUB-DMP die verschiedenen Versionen sichtbar und können als PDF exportiert, jedoch nicht mehr verändert werden.

d) Exportfunktion

Ein DMP kann als PDF exportiert werden, um beispielsweise einem Projektbericht beigelegt zu werden. Beim Export werden nur die Felder, die ausgefüllt wurden, berücksichtigt und nur deren Inhalte im PDF angezeigt. Leer gelassene Felder werden beim Export übergangen, so dass keine „Lücken“ im Text entstehen.

31 tubIT stellt einen Shibboleth-Dienst zur verteilten Authentisierung und Autorisierung für Webanwendungen und Webservices bereit: Ein Benutzer muss sich nur einmal bei seiner Heimateinrichtung (hier: TU Berlin) authentisieren, um ortsunabhängig auf Dienste oder lizenzierte Inhalte verschiedener Anbieter (hier: TUB-DMP) zugreifen zu können (Single-Sign-on).

32 European Commission, H2020, 6–9.

33 Ebd., 5.

5. Perspektive

Eines der Prinzipien des SZF ist die bedarfsbezogene Entwicklung und Weiterentwicklung benutzerorientierter Anwendungen, so auch bei TUB-DMP. Die Zusammenarbeit mit Pilotpartnern wird fortgeführt und TUB-DMP entsprechend der Anforderungen der Forschenden weiterentwickelt. Im Fokus steht derzeit die Implementierung weiterer Templates, sowohl auf die Anforderungen von Förderorganisationen bezogen als auch disziplinspezifische Templates. Auch weitere Exportmöglichkeiten (csv, XML) sind vorbereitet und können bei Bedarf zügig implementiert werden.

In Planung ist die Verlinkung zwischen TUB-DMP und DepositOnce, um die DMPs zu veröffentlichen. TUB-DMP ist inhaltlich eng an das Repositorium DepositOnce gekoppelt: In DepositOnce werden Ergebnisse aus Forschungsprojekten, in TUB-DMP die DMPs mit Informationen zu den zugehörigen Forschungsprojekten gespeichert. Die Veröffentlichung soll durch Verlinkung des DMPs in DepositOnce realisiert werden, ist somit im eigentlichen Sinne keine Funktion von TUB-DMP, sondern von DepositOnce, sei hier im Zusammenhang mit dem DMP aber trotzdem erwähnt. Die Wissenschaftlerinnen und Wissenschaftler sollen selbst entscheiden, ob sie ihre DMPs veröffentlichen, und dies auf einfache Weise tun können, indem sie die DMPs beim jeweiligen Forschungsergebnis in DepositOnce verlinken.

Literaturverzeichnis

- BMBF. „Bekanntmachung: Richtlinie zur Förderung transnationaler Verbundvorhaben auf dem Gebiet der Systemmedizin innerhalb des ERA-Netzes ERA-CoSysMed.“ *Bundesanzeiger*, 03.02.2017. Zuletzt geprüft am 15.02.2018. <https://www.bmbf.de/foerderungen/bekanntmachung-1309.html>.
- DFG. *Leitfaden für die Antragstellung: Projektanträge*. Bonn: DFG, 2018. Zuletzt geprüft am 15.02.2018. http://www.dfg.de/formulare/54_01/54_01_de.pdf.
- DFG. *Leitlinien zum Umgang mit Forschungsdaten*. Bonn: DFG, 2015. Zuletzt geprüft am 15.02.2018. http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf.
- DFG. *Umgang mit Forschungsdaten: DFG-Leitlinien zum Umgang mit Forschungsdaten*. Aktualisierungsdatum: 29.10.2015. Bonn: DFG, 2018. Zuletzt geprüft am 15.02.2018. http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/.
- Digital Curation Centre. *Checklist for a Data Management Plan: v4.0*. Edinburgh: Digital Curation Centre, 2013. Zuletzt geprüft am 15.02.2018. http://www.dcc.ac.uk/webfm_send/1279.

- Donelly, Martin und Sarah Jones. *Data Management Plan Content Checklist: Draft Template for Consultation*. Edinburgh: Digital Curation Centre, 2009. Zuletzt geprüft am 15.02.2018. http://www.dcc.ac.uk/webfm_send/1280.
- European Commission. *H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020*. Version 3.0, 26. July 2016. Brussels: European Commission, 2016. Zuletzt geprüft am 15.02.2018. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.
- Europäische Kommission. „Kommission startet Pilotprojekt zur Öffnung der mit öffentlicher Förderung erzeugten Forschungsdaten.“ *Pressemitteilung* (16. Dezember 2013). Zuletzt geprüft am 15.02.2018. http://europa.eu/rapid/press-release_IP-13-1257_de.htm.
- Fürste, Fabian. „TUB-DMP – Der Datenmanagementplan als Bindeglied zwischen Forschungsinformationssystem und Repositorium“. Vortrag auf dem 103. Deutschen Bibliothekartag in Bremen, 2014. Vortragsfolien unter <https://nbn-resolving.org/urn:nbn:de:0290-opus-17540>.
- Fürste, Fabian. „TUB-DMP – Ein Werkzeug zur Erstellung von Datenmanagementplänen an der TU Berlin.“ *DINI-Workshop „Datenmanagementpläne und ihre Bedeutung im Forschungsdatenmanagement“*. Berlin, 2015. Zuletzt geprüft am 11.04.2018. <http://www.forschungsdaten.org/images/d/dc/06--fuerst--werkzeug-erstellung-datenmanagementplaene.pdf>.
- Kuberek, Monika. *Organisatorisch-technisches Konzept für eine Forschungsdaten-Infrastruktur in der TU Berlin*. Berlin: Technische Universität, 2012. <http://dx.doi.org/10.14279/depositonce-6604>.
- Kuberek, Monika. „Der Umgang mit Forschungsdaten: Für ‚Horizon 2020‘-Anträge müssen Datenmanagementpläne entwickelt werden; Die Universitätsbibliothek hat ein neues Tool entwickelt, um Forschende zu unterstützen. *TU intern* Nr. 7–9 (2016), 4. Zuletzt geprüft am 13.04.2018. <http://archiv.pressestelle.tu-berlin.de/tui/16juli/#/4>.
- Ludwig, Jens und Harry Enke, Hrsg. *Leitfaden zum Forschungsdaten-Management – Handreichungen aus dem WissGrid-Projekt*. Glückstadt: vwh, 2013. Zuletzt geprüft am 15.02.2018. http://www.forschungsdaten.org/images/b/b0/Leitfaden_Data-Management-WissGrid.pdf.
- Wilkinson, Mark D. et al. „The FAIR Guiding Principles for scientific data management and stewardship.“ *Sci. Data* 3:160018 (2016). <https://doi.org/10.1038/sdata.2016.18>.