

# Das DARIAH-DE Repository. Elementarer Teil einer modularen Infrastruktur für geistes- und kulturwissenschaftliche Forschungsdaten

Lisa Klaffki, Herzog August Bibliothek Wolfenbüttel

Beata Mache, Niedersächsische Staats- und Universitätsbibliothek Göttingen

## Zusammenfassung:

DARIAH-DE unterstützt mit digitalen Ressourcen und Methoden arbeitende Geistes- und Kulturwissenschaftler/innen in Forschung und Lehre. Forschungsdaten und Wissenschaftliche Sammlungen haben dabei eine zentrale Bedeutung. Im Rahmen einer *Forschungsdaten-Föderationsarchitektur* stehen eine Reihe von Werkzeugen zur Verfügung, mit denen Daten gefunden, vernetzt, publiziert und archiviert werden können. Hierzu zählt auch das *DARIAH-DE Repository*, das den Forschenden eine sichere, langfristige und nachhaltige Speicherung sowie die Veröffentlichung der Forschungsdaten ermöglicht.

## Summary:

DARIAH-DE supports research in the humanities and cultural sciences with digital methods and procedures. One of the main pillars of DARIAH-DE is dedicated to the processing and storing of research data, for which several tools and services are provided. The *DARIAH-DE Repository* is one component of the *Data Federation Architecture* (DFA) offered by DARIAH-DE to manage research data. The repository is a digital long-term archive for research data enabling researchers to store and publish data in a secure and sustainable manner.

**Zitierfähiger Link (DOI):** <https://doi.org/10.5282/o-bib/2018H3S92-103>

**Autorenidentifikation:** Lisa Klaffki, ORCID: <https://orcid.org/0000-0002-8791-7422>; Beata Mache, ORCID: <https://orcid.org/0000-0001-5251-081X>

**Schlagwörter:** Forschungsdatenmanagement, Repositorien, DARIAH-DE, Forschungsdateninfrastruktur

## 1. Einleitung

DARIAH-DE<sup>1</sup> ist ein vom Bundesministerium für Bildung und Forschung (BMBF) seit 2011 gefördertes Vorhaben, das eine Forschungsinfrastruktur für die mit digitalen Methoden arbeitenden Geistes- und Kulturwissenschaftler/inn/en entwickelt. Die bisher aufgebaute Infrastruktur, die Werkzeuge und Dienste zur Analyse, Modellierung und Speicherung der Forschungsdaten umfasst, befindet sich im Übergang zur Betriebsphase. Schon jetzt wird sie von mehr als 3.500 Wissenschaftler/inn/en und über 80 Forschungsvorhaben (Stand Frühjahr 2018) genutzt.

1 Portal des DARIAH-DE-Projekts, <<https://de.dariah.eu/>>, Stand: 18.02.2018.

Von Anfang an auf dem Grundsatz der „architecture of participation“ aufbauend, also eine Kultur der Teilhabe pflegend, orientiert sich die von den DARIAH-DE Partnern<sup>2</sup> entwickelte digitale Forschungsinfrastruktur an den Bedarfen der Nutzenden und ermöglicht ihr gleichzeitig, in anderen Kontexten entstandene Dienste, Werkzeuge, fachwissenschaftliche Inhalte, Daten oder Lehrkonzepte allen Interessierten zugänglich zu machen.

Thematisch fokussiert das DARIAH-DE Vorhaben einerseits geistes- und kulturwissenschaftliche Schwerpunkte wie „Quantitative Textanalyse“ oder „Analysieren, Annotieren, Visualisieren“, andererseits die Evaluation des Einflusses der digitalen Werkzeuge und Daten auf die Forschungspraxis.<sup>3</sup>

Neben den geistes- und kulturwissenschaftlich orientierten Schwerpunkten sind Forschungsdaten und Wissenschaftliche Sammlungen für DARIAH-DE von zentraler Bedeutung. So stehen im modularen System der DARIAH-DE Forschungsdaten-Föderationsarchitektur<sup>4</sup> den Forschenden Werkzeuge zur Verfügung, mit denen sie ihre Daten beschreiben, verwalten, archivieren, publizieren und sie mit den Daten der anderen Forschenden vernetzen können, so dass sie gefunden und nachgenutzt werden.

Die technische Säule der DARIAH-DE Forschungsinfrastruktur ermöglicht eine sichere, nachhaltig verfügbare und persistente Datenspeicherung. Support für die Nutzerschaft leistet ein Helpdesk-System.<sup>5</sup>

## 2. Der Research Data Lifecycle als Konzept

Die Werkzeuge und Dienste wurden nicht nur in Zusammenarbeit mit den potentiellen Nutzenden entwickelt, sondern basieren auch auf konzeptioneller Vorarbeit. Neben Definitionen zu zentralen Begriffen wie „Forschungsdaten“ und „Wissenschaftlichen Sammlungen“<sup>6</sup> spielte dabei der sogenannte Research Data Lifecycle eine wichtige Rolle: Am Ende des Forschungsprozesses steht traditionell die Veröffentlichung der Ergebnisse in Form eines Artikels oder einer Monographie. In einem Research Data Lifecycle – der Name deutet es schon an – gibt es kein Ende im eigentlichen Sinne. Zwar findet das Speichern und Veröffentlichen von Daten auch hier am Ende des Forschungsprozesses statt, die zugänglich gemachten Daten können dann aber die Quelle einer neuen Forschungsfrage werden.

Um die Schritte zu modellieren, die die Forschungsdaten dabei durchlaufen und um zu verdeutlichen, welche Anforderungen dabei an die Daten und die Infrastruktur gestellt werden, lässt sich ein

2 Forschungsverbund, DARIAH-DE, <<https://de.dariah.eu/der-forschungsverbund>>, Stand: 18.02.2018.

3 Die Impactomatrix, DARIAH-DE, <<https://dariah-de.github.io/Impactomatrix/>>, Stand: 18.02.2018.

4 Die Data Federation Architecture, DARIAH-DE, <<https://de.dariah.eu/data-federation-architecture>>, Stand: 18.02.2018; Gradl, Tobias; Henrich, Andreas: Die DARIAH-DE-Föderationsarchitektur. Datenintegration im Spannungsfeld forschungsspezifischer und domänenübergreifender Anforderungen, in: Bibliothek Forschung und Praxis 40 (2), 2016, S. 222-228. Online: <<http://dx.doi.org/10.1515/bfp-2016-0027>>; Gradl, Tobias; Henrich, Andreas; Plutte, Christoph: Heterogene Daten in den Digital Humanities: Eine Architektur zur forschungsorientierten Föderation von Kollektionen, in: Baum, Constanze; Stäcker, Thomas (Hg.): Grenzen und Möglichkeiten der Digital Humanities. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1. Wolfenbüttel 2015. Online: <[http://nbn-resolving.de/10.17175/sb001\\_020](http://nbn-resolving.de/10.17175/sb001_020)>.

5 Helpdesk, DARIAH-DE, <<https://de.dariah.eu/basisinfrastruktur-helpdesk>>, Stand: 18.02.2018.

6 Oltersdorf, Jenny; Schmunk, Stefan: Von Forschungsdaten und wissenschaftlichen Sammlungen. Zur Arbeit des Stakeholdergremiums ‚Wissenschaftliche Sammlungen‘ in DARIAH-DE, in: Bibliothek Forschung und Praxis 40 (2), 2016, S. 179-185. Online: <<http://dx.doi.org/10.1515/bfp-2016-0036>>.

Research Data Lifecycle aufstellen. Dieses Konzept aus der Informationswissenschaft wird in der einen oder anderen Abwandlung in vielen Forschungsinfrastrukturen und Projekten genutzt.<sup>7</sup> Es gibt kein verbindliches, einheitliches Modell. Die einzelnen Schemata spiegeln heterogene wissenschaftliche Praktiken und (Daten-)Standards, vor allem aber verschiedene Sichtweisen auf Daten, sei es die Perspektive der Forschenden oder jener für das Datenmanagement Zuständigen.<sup>8</sup>

Innerhalb von DARIAH-DE wurde deshalb ein Modell entwickelt, das disziplinenübergreifend in den digitalen Geisteswissenschaften anwendbar ist und damit als intellektuelles Konzept den theoretischen Unterbau der Forschungsinfrastruktur bildet.<sup>9</sup> Der generische Workflow umfasst typische Arbeitsschritte im geisteswissenschaftlichen Forschungsprozess<sup>10</sup> und wird iterativ zu einem Research Data Lifecycle erweitert, um neue Forschungsprozesse einbinden und Zwischenergebnisse oder Forschungsdaten, etwa in einem Repositorium, publizieren zu können.<sup>11</sup>

### 3. Die DARIAH-DE Data Federation Architecture als Proof of Concept

Unter dem Begriff der *Data Federation Architecture* (DFA) sind mehrere Komponenten gebündelt, von denen jede für sich alleine oder im Zusammenspiel genutzt werden kann:<sup>12</sup> Das DARIAH-DE Repository,<sup>13</sup> die *Collection Registry*, das *Data Modeling Environment*<sup>14</sup> und die *Generic Search* sowie ein EPIC-PID-Service.<sup>15</sup> Dabei geht es nicht um eine Integration, also „Gleichmachung“ von heterogenen Daten, sondern darum, eine Arbeitsumgebung bereitzustellen, in der Daten aus verschiedenen Quellen gemeinsam vorgehalten und durchsucht werden können, also in einer Föderation zusammengestellt sind.

7 Für ein vereinfachtes, in Anlehnung an das DCC Curation-Lifecycle-Model (vgl. Anm. 8) entwickeltes Modell s. Ludwig, Jens; Enke, Harry (Hg.): Leitfaden zum Forschungsdaten-Management. Handreichungen aus dem WissGrid-Projekt. Glückstadt 2013, S. 14–16; IANUS, IT-Empfehlungen in den Altertumswissenschaften, <<https://www.ianus-fdz.de/it-empfehlungen/lebenszyklus>>, Stand: 16.2.2018; UK Data Archive, <<http://www.data-archive.ac.uk/create-manage/life-cycle>>, Stand: 16.02.2018.

8 Vgl. das Curation-Lifecycle-Model und das *Data Curation Continuum*, Rümpel, Stefanie: Der Lebenszyklus von Forschungsdaten, in: Büttner, Stephan; Hobohm, Hans-Christoph; Müller, Lars (Hg.): Handbuch Forschungsdatenmanagement, Bad Honnef: 2011. S. 25–34. Online: <<http://nbn-resolving.de/urn:nbn:de:kobv:525-opus-2412>>.

9 Puhl, Johanna; Andorfer, Peter; Höckendorff, Mareike; Schmunk, Stefan; Stiller, Juliane; Thoden, Klaus: Diskussion und Definition eines Research Data LifeCycle für die digitalen Geisteswissenschaften. DARIAH-DE Working Papers Nr. 11. Göttingen 2015. Online: <<http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-4-4>>.

10 Ebd., S. 41.

11 Ebd., S. 43.

12 Gradl; Henrich; Plutte: Heterogene Daten, 2015, Abb. 1.

13 Ausführliche Beschreibung folgt im nächsten Kapitel.

14 In älterer Literatur, Präsentationen etc. getrennt unter „Schema Registry“ und „Crosswalk Registry“ vorgestellt, wurden die beiden Komponenten im Zuge ihrer Weiterentwicklung im Sommer 2017 zusammengelegt. Die Dokumentation ist zur Zeit noch in Bearbeitung, sodass die Begriffe parallel zu finden sind.

15 Nachhaltige Referenzierung mithilfe persistenter Identifikatoren, DARIAH-DE, <<https://de.dariah.eu/persistent-identifiers>>, Stand: 16.02.2018.

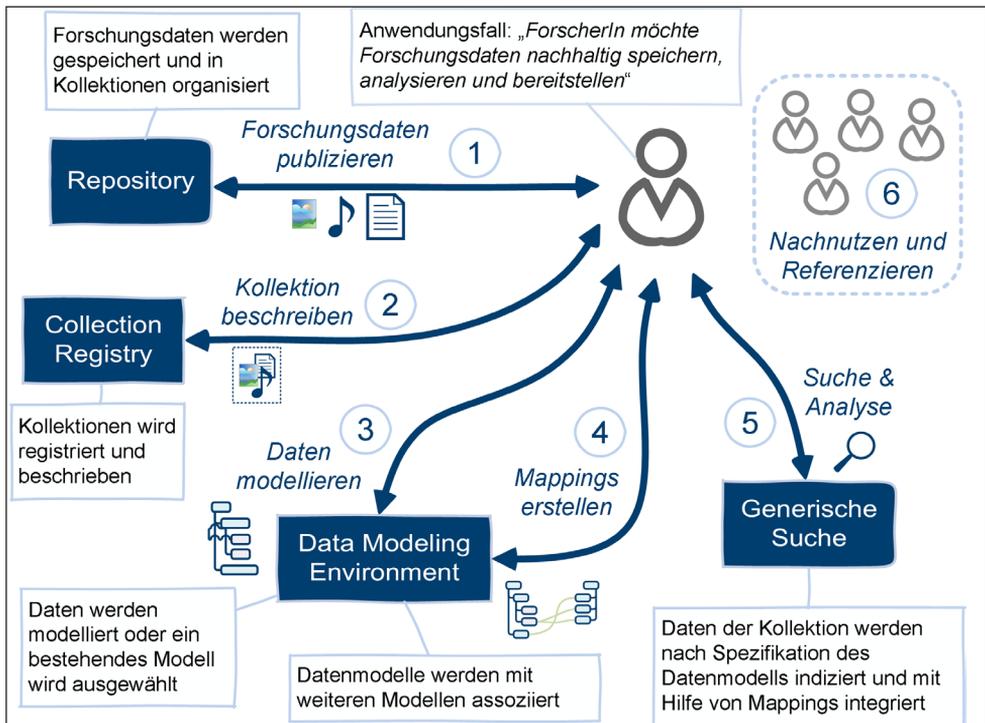


Abb. 1: Schematischer Aufbau der DARIAH-DE Data Federation Architecture.<sup>16</sup>

Die Komponenten der *Data Federation Architecture* können die Forschenden bei vielen Aufgaben des Forschungsdatenzyklus unterstützen: In der *Generischen Suche* kann man über heterogene Datenbestände hinweg nach für die eigene Forschungsfrage passenden Daten suchen. Forschungsdaten, die unterschiedliche Schemata aufweisen, können mit dem *Data Modeling Environment* gemappt werden. Dabei können bereits hinterlegte Metadatenstandards und Crosswalks zwischen Schemata nachgenutzt oder neue Modelle und Mappings angelegt werden. Aufgrund der Komplexität ist diese Komponente eher als Expertentool anzusehen und nur mit Unterstützung aus dem Projektkontext heraus und / oder entsprechender Einarbeitung nutzbar. Perspektivisch wird aus den umgesetzten Use Cases eine „Grammatik“ für die Mapping-Sprache erstellt, die auch dieses Tool künftig einer breiteren Nutzerschaft zugänglich machen wird.

Analyse und Visualisierung der Ergebnisse werden nicht durch die DFA abgedeckt, da diese Arbeitsschritte sehr spezifische Tools erfordern, die sich nicht durch eine generische Lösung abdecken lassen.

<sup>16</sup> Gradl, Tobias: Schematischer Aufbau der DARIAH-DE Data Federation Architecture, DARIAH-DE, <<https://de.dariah.eu/data-federation-architecture>>, Stand: 30.06.2018.

In den digitalen Geisteswissenschaften stehen hinreichend viele Tools zur Verfügung, die diese Lücke zielgerichtet füllen können.<sup>17</sup>

Ergebnisse und im Forschungsprozess entstandene Daten können dann im *DARIAH-DE Repository* mit einem persistenten Identifier veröffentlicht und archiviert werden. Solche Forschungsdatensammlungen bzw. ihre Metadaten können durch automatische Übernahme aus dem Repositorium und, optional, durch Erweiterung der Beschreibung anhand des DARIAH Collection Data Description Model (DCDDM)<sup>18</sup> in der *Collection Registry* verzeichnet werden.

Die über das *DARIAH-DE Repository* publizierten Sammlungen weisen automatisch eine maschinenlesbare Schnittstelle zu den beschriebenen Daten auf. Dadurch können sie wiederum in der – am Anfang des hier skizzierten Forschungsprozesses stehenden – *Generischen Suche* indiziert und fortan gefunden werden. Aber auch unabhängig vom *DARIAH-DE Repository* können andernorts archivierte Datensammlungen in der *Collection Registry* beschrieben werden. Diese werden – sofern eine Schnittstelle für den Datenzugriff zur Verfügung steht – ebenfalls in der *Generischen Suche* indiziert; andernfalls sind sie bisher nur in der *Collection Registry* auffindbar. Für die Zukunft ist geplant, via OAI-PMH-Schnittstelle auch die Beschreibungen selbst über die *Generische Suche* zugänglich zu machen.

Die Wissenschaftler/innen können also an derjenigen Stelle in den Forschungsdatenzzyklus bzw. die DFA „einsteigen“, die sie für ihren Forschungsprozess benötigen. Die beiden Konzepte – theoretischer Unterbau und praktische Umsetzung – können dabei nicht „Eins zu Eins“ aufeinander bezogen werden, sie sind vielmehr durch „one-to-many“-Beziehungen miteinander verbunden.

## 4. DARIAH-DE Publikator und Repository

Das im Dezember 2017 in den produktiven Betrieb überführte *DARIAH-DE Repository* steht nicht nur Forschenden, sondern auch Gedächtniseinrichtungen als digitaler Speicher für geistes- und kulturwissenschaftliche Forschungsdaten zur Verfügung. Wie alle anderen DARIAH-DE Dienste und Werkzeuge<sup>19</sup> ist das Repositorium an keinen Standort gebunden, sondern ein fachspezifisches, überinstitutionelles Angebot<sup>20</sup> und kann somit potenziell die für die Geistes- und Kulturwissenschaften relevanten Forschungsdaten unabhängig von ihrer Herkunft speichern. Diese Lösung ist gerade für Verbund- oder multinationale Projekte sowie für außeruniversitäre Forschungseinrichtungen, die keine eigene digitale Forschungsdateninfrastruktur aufbauen können oder wollen, besonders wichtig. Auch Gedächtniseinrichtungen wie Archive, Museen und Bibliotheken können auf diese Weise

---

17 Eine aktuelle Abfrage des DiRT-Verzeichnisses liefert beispielsweise 118 Ergebnisse für „Visualization“ (<<http://dirtdirectory.org/tadirah/visualization>>, Stand: 16.02.2018).

18 DARIAH Collection Data Description Model, GitHub <<https://github.com/DARIAH-DE/DCDDM>>, Stand: 16.02.2018.

19 DARIAH-DE Dienste und Werkzeuge, DARIAH-DE, <<https://de.dariah.eu/list-services>>, Stand: 16.02.2018.

20 Die Forschungsdatenrichtlinien zahlreicher Universitäten empfehlen mittlerweile die Speicherung der Forschungsdaten in – soweit vorhanden – fachspezifischen Repositorien. Vgl. dazu: Forschungsdaten-Leitlinie der Universität Göttingen (einschl. UMG) vom 01.06.2014, <<https://www.uni-goettingen.de/de/488918.html>>, Stand: 16.02.2018; vgl. auch: Richtlinie zum Umgang mit Forschungsdaten an der Leibniz Universität Hannover, 26.04.2017, <[https://www.fdm.uni-hannover.de/fileadmin/fdm/Richtlinie\\_Forschungsdaten.pdf](https://www.fdm.uni-hannover.de/fileadmin/fdm/Richtlinie_Forschungsdaten.pdf)>, Stand: 18.02.2018.

ihre digitalen Objekte, die für die Geistes- und Kulturwissenschaften grundlegende, die Bearbeitung ihrer Forschungsfragen tatsächlich erst ermöglichende Forschungsdaten sind, sichtbar machen und dadurch ihre Nachnutzung stärken.<sup>21</sup>

Auch deshalb folgt DARIAH-DE einem sehr weiten Verständnis von Forschungsdaten:

„Unter digitalen geistes- und kulturwissenschaftlichen Forschungsdaten werden innerhalb von DARIAH-DE all jene Quellen / Materialien und Ergebnisse verstanden, die im Kontext einer geistes- und kulturwissenschaftlichen Forschungsfrage gesammelt, erzeugt, beschrieben und / oder ausgewertet werden und in maschinenlesbarer Form zum Zwecke der Archivierung, Zitierbarkeit und zur weiteren Verarbeitung aufbewahrt werden können.“<sup>22</sup>

Diese – die Spezifik der geistes- und kulturwissenschaftlichen Forschung berücksichtigende – Auffassung stellt die Forschungsinfrastruktur vor technische, fachwissenschaftliche, rechtliche und förderpolitische Herausforderungen. Während das seit 2015 von DARIAH-DE betreute *TextGrid Repository*<sup>23</sup> allein Daten im XML/TEI-Format hostet, ermöglicht das *DARIAH-DE Repository*<sup>24</sup> die Speicherung von Daten in allen gängigen Formaten. Vor der Veröffentlichung können die Daten in einem eigenen Speicherbereich (*OwnStorage*<sup>25</sup>) in Kollektionen und Subkollektionen gespeichert, versioniert, verwaltet und – wenn gewünscht – gelöscht werden.

Damit die Forschenden auch ohne weitere fachliche Hilfe (etwa aus dem Bibliotheksbereich) ihre Daten selbst verwalten und publizieren können, erfolgt die Metadatenerfassung entsprechend dem Dublin Core Simple Standard.<sup>26</sup> Verpflichtend sind nur Metadatenangaben zu Titel, Autor/in und der Lizenzierung, alle weiteren Felder sind optional. Die Nutzer/innen erhalten Erklärungen zu den einzelnen Feldern und werden angehalten, diese auch auszufüllen und detailliertere Beschreibungen vorzunehmen – dies geschieht aber freiwillig (Vgl. Abb. 4).

---

21 Die Oberfläche des *DARIAH-DE Publikators* ist in deutscher, englischer, französischer, niederländischer und polnischer Sprache verfügbar. Dieses Angebot richtet sich insbesondere an Gedächtniseinrichtungen und multinationale Forschungsgruppen.

22 Oltersdorf und Schmunk: Forschungsdaten, 2016, S. 181.

23 TextGrid, <<https://textgridrep.org/>>, Stand: 18.02.2018; vgl. auch: User Guide, Das DARIAH-DE Repository und das TextGrid Repository, 14.02.2018, <<https://wiki.de.dariah.eu/display/publicde/Das+DARIAH-DE+Repository+und+das+TextGrid+Repository>>, Stand: 18.02.2018.

24 DARIAH-DE Repository, DARIAH-DE, <<https://de.dariah.eu/repository>>, Stand: 18.02.2018

25 DARIAH-DE Publikator, DARIAH-DE, <<https://de.dariah.eu/publikator>>, Stand: 18.02.2018.

26 Die fachspezifischen Metadaten-Schemata werden im Data Modeling Environment miteinander assoziiert. Vgl. Abb. 3. Die Einbindung weitere Metadaten-Schemata in den DARIAH-DE Publikator ist geplant.

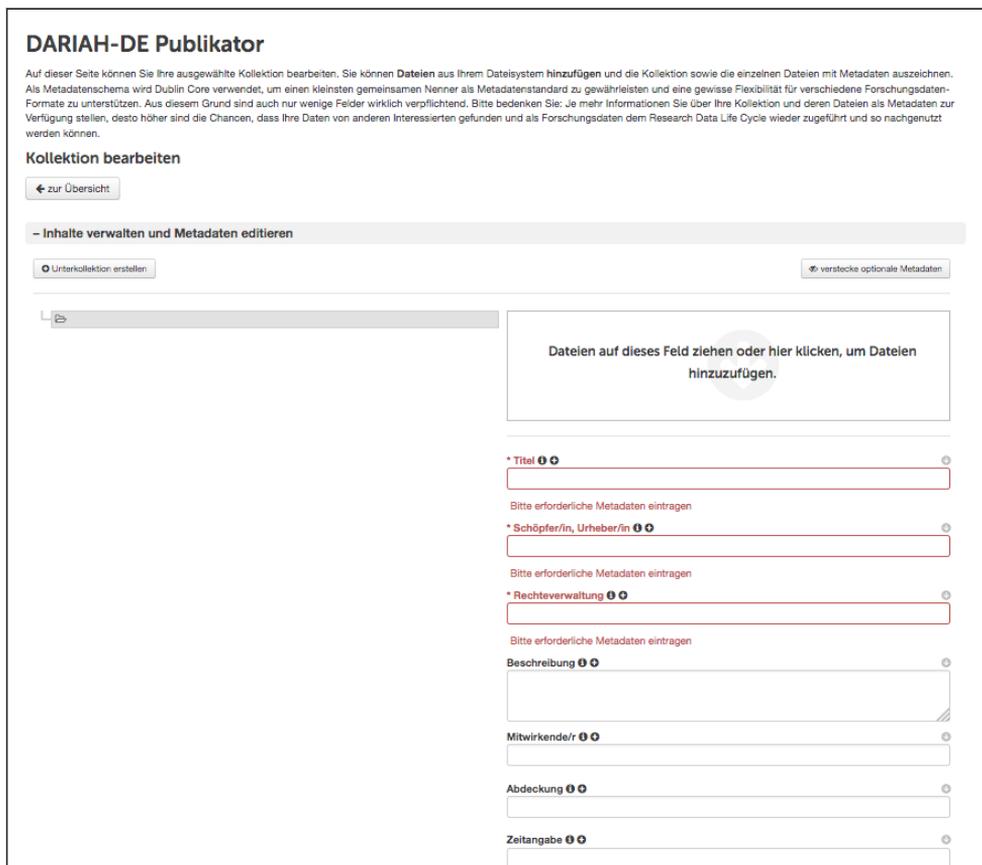


Abb. 2: Metadaten-Pflichtfelder im DARIAH-DE Publikator.<sup>27</sup>

In den Kollektionen können beliebig viele Subkollektionen eingerichtet werden, die wiederum beliebig viele Datensätze enthalten können.<sup>28</sup> Die Datenveröffentlichung ist ein durch den oder die Datenhalter/in durchzuführender Akt. Im Prozess der Veröffentlichung werden die Daten – für die Nutzer/innen unsichtbar – aus dem *OwnStorage* des/der Nutzenden in den *PublicStorage* (Abb. 4) verschoben.

27 DARIAH-DE Publikator, <<https://repository.de.dariah.eu/publikator/>>, Stand: 30.06.2018.

28 Terms of Use, 01.12.2017, <<https://repository.de.dariah.eu/1.0/dhcrud/21.11113/0000-000B-C8F1-3/data>>, Stand: 18.02.2018.

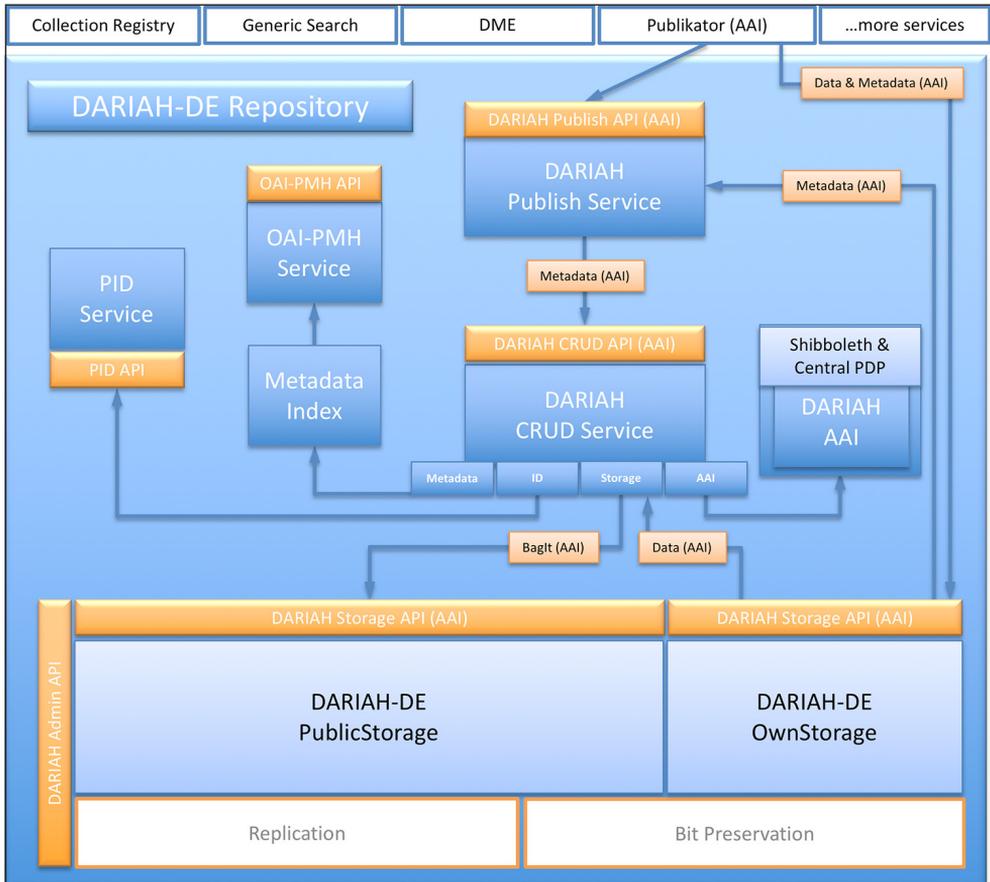


Abb. 3: DARIAH-DE Repository: technischer Aufbau.<sup>29</sup>

Jedes Datum und jede (Sub-)Kollektion wird mit einem Persistenten Identifier (DOI)<sup>30</sup> ausgezeichnet sowie mit deskriptiven, technischen und administrativen Metadaten ergänzt (Abb. 5). Ob die so veröffentlichten Daten nur gespeichert und dauerhaft referenziert oder auch in der *Generischen Suche* von DARIAH-DE (Abb. 2) indiziert werden sollen, kann der/die Veröffentliche selbst entscheiden, indem er/sie der Aufnahme der Metadaten in die *DARIAH-DE Collection Registry* (mittels OAI-PMH Schnittstelle) zustimmt (Abb. 6). Über die Auflösung von DOI werden die veröffentlichten Forschungsdaten gefunden und sind in gängigen Katalogen und Suchmaschinen zu finden.<sup>31</sup>

29 Funk, Stefan: Aufbau des DARIAH-DE Repositorys, <<https://de.dariah.eu/repository>>, Stand: 30.06.2018.

30 Digital Object Identifier (DOI), <<https://www.doi.org/>>, Stand: 18.02.2018.

31 Zu nennen wären z.B.: DataCite, <<https://www.datacite.org/>>, Stand: 18.02.2018 und BASE, <<https://www.base-search.net/about/de/>>, Stand: 18.02.2018.

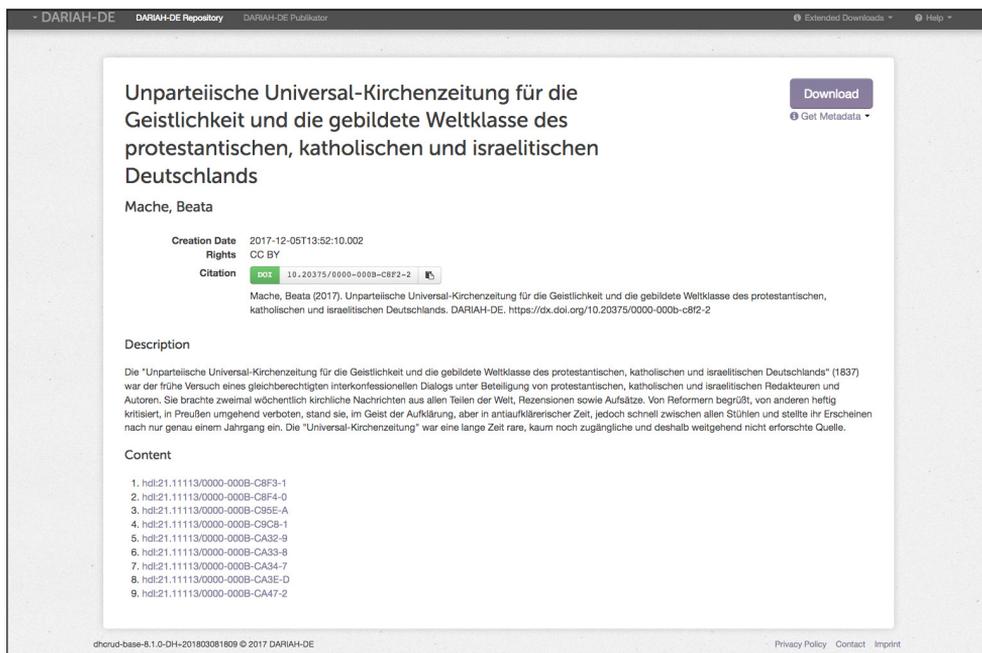


Abb. 4: Metadaten im DARIAH-DE Repository.

Dieser aus wenigen intuitiven Schritten bestehende Workflow – Daten hochladen, sie in Subkollektionen organisieren, mit Metadaten beschreiben, Lizenzen bestimmen, Terms of Use<sup>32</sup> akzeptieren – zeigt, dass sich das DARIAH-DE Repository als ein niedrigschwelliges Angebot direkt an die Datenhaltenden und -erzeugenden richtet, die eigenverantwortlich und selbstbestimmt, aber den FAIR-Prinzipien entsprechend,<sup>33</sup> ihre Daten veröffentlichen können.

Den dauerhaften und zuverlässigen Betrieb des DARIAH-DE Repositorys sichert das Humanities Data Centre an der Universität Göttingen.<sup>34</sup>

32 DARIAH-DE Repository Terms of Use, 24.04.2018, <<https://dx.doi.org/10.20375/0000-000B-CB47-1>>, Stand: 24.04.2018.

33 The FAIR Data Principles, Force11, <<https://www.force11.org/group/fairgroup/fairprinciples>>, Stand: 18.02.2018.

34 Service Portfolio, Humanities Data Centre, <<https://humanities-data-centre.de/>>, Stand: 18.02.2018.

The screenshot shows the DARIAH-DE Publikator interface. At the top, there is a navigation bar with the DARIAH-DE logo and the title 'DARIAH-DE Publikator'. Below this, a heading reads 'DARIAH-DE Publikator'. A paragraph explains that users can create new collections and manage existing ones. A section titled 'Übersicht über Ihre Kollektionen' contains a button for 'neue Kollektion anlegen'. A collection entry is displayed with the following details:

- Titel:** Unparteiische Universal-Kirchenzeitung für die Geistlichkeit und die gebildete Weltklasse des protestantischen, katholischen und israelitischen Deutschlands
- Beschreibung:** Die "Unparteiische Universal-Kirchenzeitung für die Geistlichkeit und die gebildete Weltklasse des protestantischen, katholischen und israelitischen Deutschlands" (1837) war der frühe Versuch eines gleichberechtigten interkonfessionellen Dialogs unter Beteiligung von protestantischen, katholischen und israelitischen Redakteuren und Autoren. Sie brachte zweimal wöchentlich kirchliche Nachrichten aus allen Teilen der Welt, Rezensionen sowie Aufsätze. Von Reformern begrüßt, von anderen heftig kritisiert, in Preußen umgehend verboten, stand sie, im Geist der Aufklärung, aber in antiaufklärerischer Zeit, jedoch schnell zwischen allen Stühlen und stellte ihr Erscheinen nach nur genau einem Jahrgang ein. Die "Universal-Kirchenzeitung" war eine lange Zeit rare, kaum noch zugängliche und deshalb weitgehend nicht erforschte Quelle.
- Schöpfer:** Mache, Beata
- Status:** registriert
- Anzahl der Kollektionen:** 6
- Anzahl der Dateien:** 336
- DOI ihrer Kollektion:** 10.20375/0000-000B-CBF2-2

Below the details, a text block reads: 'Herzlichen Glückwunsch! Zusätzlich zur Publikation im DARIAH-DE Repository und der Referenzierbarkeit über den angezeigten Identifikator (DOI) ist Ihre Kollektion nun auch als Sammlung in der Collection Registry registriert und wird nun auch durch die Generische Suche indiziert. Ihre Inhalte sind dort nun öffentlich recherchierbar.' At the bottom, there are buttons for 'Collection Registry', 'Generische Suche', 'Im Repository anzeigen', and 'Kollektion löschen'.

Abb. 5: Eine veröffentlichte Kollektion mit der Option der Aufnahme in die DARIAH-DE Collection Registry.

Die beiden von DARIAH-DE betriebenen Repositorien erfüllen die Kriterien der stabilen Referenzierbarkeit, Nachprüfbarkeit und Nachvollziehbarkeit sowie die ethischen und rechtlichen Anforderungen, weswegen sie 2018 mit dem CoreTrustSeal<sup>35</sup> zertifiziert werden: zuerst das *TextGrid Repository*, anschließend das *DARIAH-DE Repository*. Damit soll die Vertrauenswürdigkeit der beiden Repositorien verdeutlicht werden, sowohl für die nach Daten suchende Nutzerschaft als auch für die ihre Daten und Ergebnisse publizierenden Wissenschaftler/innen, die durch die Nutzung zertifizierter Dienste den Anforderungen der Förderer an ein Forschungsdatenmanagement zuverlässig nachkommen können.

## 5. Durch Modularität fit für die Zukunft

Ein Resümee oder eine Übersicht zur Nutzung des *DARIAH-DE Repositorys* ist wenige Monate nach der Bereitstellung noch nicht möglich. Damit es aber dem Bedarf der Nutzerschaft von Beginn an entgegenkommt, wurde diese in der Entwicklungsphase eingebunden, u.a. im Rahmen von Experten-Workshops.<sup>36</sup> Einige der Empfehlungen konnten so schon vor dem Übergang in die Produktivphase berücksichtigt werden. Zu den aktuell geplanten Funktionserweiterung gehört die Einführung eines Gruppenmodus, damit auch vor der Veröffentlichung Teams kollaborativ im *Own Storage* Forschungsdaten verwalten können.<sup>37</sup> Eine Herausforderung ist die aus der Gemeinschaft der Forschenden

35 Why Certification, CoreTrustSeal, <<https://www.coretrustseal.org/why-certification/>>, Stand: 18.02.2018.

36 Siehe u.a. Workshop „Daten- und Metadatenformate und Import in das DARIAH-DE Repository“, DARIAH-DE, 15./16. August 2017, <<https://de.dariah.eu/2017-08-15-repository-workshop-goettingen>>, Stand: 18.02.2018.

37 Die Benennung mehrerer Autor/inn/en bei der Veröffentlichung der Daten ist selbstverständlich auch schon in der aktuellen Version möglich.

kommende Anforderung, schon forschungsbegleitend Daten persistent referenziert veröffentlichen zu können. Dies ist zurzeit zwar möglich, führt aber zur Teilung der zusammengehörenden Daten in viele – oft sehr kleine – Kollektionen, da einmal veröffentlichte Datensätze nicht erweitert werden können.

Die modulare Struktur der DFA erlaubt eine solche Weiterentwicklung und damit eine Anpassung an sich verändernde Nutzungsszenarien und neue Entwicklungen im Forschungsdatenmanagement. Einzelne Komponenten können unabhängig voneinander modifiziert werden. Ebenso ist es möglich, nur einige Module in eine eigene Infrastruktur zu integrieren. So nutzt der Forschungsverbund Marbach Weimar Wolfenbüttel eigene Instanzen der *Collection Registry*, des *Data Modeling Environment* und der *Generischen Suche* für dessen *Virtuellen Forschungsraum*,<sup>38</sup> einer virtuellen Forschungsumgebung, unabhängig von der DARIAH-DE Instanz und modifiziert sie den eigenen Ansprüchen entsprechend. Die innerhalb dieses Use Cases stattfindenden Entwicklungen werden, soweit sinnvoll, in die generische DARIAH-DE Instanz zurückgespielt.<sup>39</sup>

DARIAH-DE bemüht sich, den Nutzenden ausreichende Unterstützung bei der Verwendung der Dienste zu bieten, individuelle Betreuung ist aber nur begrenzt möglich. Deshalb werden, gerade auch für Einsteiger/innen, fortlaufend ausführliche Dokumentationen<sup>40</sup> und Schulungsmaterialien – sowohl auf der DARIAH-DE Webseite, als auch in dem in der Community populären DHd-Blog – bereitgestellt.<sup>41</sup>

## Literaturverzeichnis

- Gradl, Tobias; Aschauer, Anna; Dogunke, Swantje; Klaffki, Lisa; Schmunk, Stefan; Steyer, Timo: Daten sammeln, modellieren und durchsuchen mit DARIAH-DE, 2017. <<http://dx.doi.org/10.5281/zenodo.582316>>.
- Gradl, Tobias; Henrich, Andreas: Die DARIAH-DE-Föderationsarchitektur. Datenintegration im Spannungsfeld forschungsspezifischer und domänenübergreifender Anforderungen, in: *Bibliothek Forschung und Praxis* 40 (2), 2016, S. 222–228. Online: <<http://dx.doi.org/10.1515/bfp-2016-0027>>.
- Gradl, Tobias; Henrich, Andreas; Plutte, Christoph: Heterogene Daten in den Digital Humanities: Eine Architektur zur forschungsorientierten Föderation von Kollektionen, in: Baum, Constanze; Stäcker, Thomas (Hg.): *Grenzen und Möglichkeiten der Digital Humanities*.

38 Virtueller Forschungsraum, <<https://vfr.mww-forschung.de/>>, Stand: 18.02.2018.

39 Gradl et al.: Daten sammeln, 2017, S. 7f.

40 Ausführliche Dokumentation und Beschreibung des Veröffentlichungsprozesses, DARIAH-DE Public Wiki, <<https://wiki.de.dariah.eu/display/publicde/Das+DARIAH-DE+Repository>>, Stand: 18.02.2018; technische Dokumentation, <<http://repository.de.dariah.eu/doc/services/>>, Stand: 18.02.2018.

41 Leone, Claudio: DARIAH-DE. Forschen mit digitalen Diensten und Werkzeugen für alle Geistes- und Kulturwissenschaftler, DHd-Blog, 13. Februar 2018. <<http://dhd-blog.org/?p=9082>>, Stand: 18.01.2018.; Mache, Beata: Meine erste Kollektion im DARIAH-DE Repository, DHd-Blog, 7. Dezember 2017. <<http://dhd-blog.org/?p=8798>>, Stand: 18.01.2018.

Sonderband der Zeitschrift für digitale Geisteswissenschaften 1. Wolfenbüttel 2015.  
Online: <[http://nbn-resolving.de/10.17175/sb001\\_020](http://nbn-resolving.de/10.17175/sb001_020)>.

- Leone, Claudio: DARIAH-DE. Forschen mit digitalen Diensten und Werkzeugen für alle Geistes- und Kulturwissenschaftler, DHd-Blog, 13. Februar 2018. <<http://dhd-blog.org/?p=9082>>, Stand: 18.01.2018.
- Ludwig, Jens; Enke, Harry (Hg.): Leitfaden zum Forschungsdaten-Management. Handreichungen aus dem WissGrid-Projekt. Glückstadt 2013.
- Mache, Beata: Meine erste Kollektion im DARIAH-DE Repository, DHd-Blog, 7. Dezember 2017. <<http://dhd-blog.org/?p=8798>>, Stand: 18.01.2018.
- Oltersdorf, Jenny; Schmunk, Stefan: Von Forschungsdaten und wissenschaftlichen Sammlungen. Zur Arbeit des Stakeholdergremiums ‚Wissenschaftliche Sammlungen‘ in DARIAH-DE, in: Bibliothek Forschung und Praxis 40 (2), 2016, S. 179–185. Online: <<http://dx.doi.org/10.1515/bfp-2016-0036>>.
- Puhl, Johanna; Andorfer, Peter; Höckendorff, Mareike; Schmunk, Stefan; Stiller, Juliane; Thoden, Klaus: Diskussion und Definition eines Research Data LifeCycle für die digitalen Geisteswissenschaften. DARIAH-DE Working Papers Nr. 11. Göttingen 2015. Online: <<http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-4-4>>.
- Rümpel, Stefanie: Der Lebenszyklus von Forschungsdaten, in: Büttner, Stephan; Hobohm, Hans-Christoph; Müller, Lars (Hg.): Handbuch Forschungsdatenmanagement, Bad Honnef: 2011. S. 25–34. Online: <<http://nbn-resolving.de/urn:nbn:de:kobv:525-opus-2412>>.