



Fachhochschule Potsdam
University of
Applied Sciences

Vergleich und Analyse von Forschungsdatenrepositorien

**Exemplarische Untersuchung des Umgangs mit
Forschungsdaten unter besonderer Betrachtung der
Ressource Video**

Bachelorarbeit

zur Erlangung des akademischen Grades

Bachelor of Arts (B.A.)

im Fachgebiet Informationswissenschaften

vorgelegt von: Melanie Janßen

Studienbereich: Bibliotheksmanagement

Matrikelnummer: 13741

Erstgutachterin: Prof. Dr. Heike Neuroth

Zweitgutachter: Prof. Dr. Günther Neher

Berlin, den 29. Januar 2018

Kurzfassung

Ziel dieser Arbeit war es, Repositorien zu analysieren, die Videos als Forschungsdaten vorhalten und im Rahmen eines Vergleiches die Gemeinsamkeiten und Unterschiede von verschiedenen Aspekten herauszufinden. Eine Auswahl von acht Repositorien wurde qualitativ analysiert. Die Erhebung der Daten erfolgte durch die inhaltliche Analyse der Webseiten der Repositorien anhand eines nachgenutzten Kriterienkataloges. Die Ergebnisse der Analyse wurden anhand der Kriterien miteinander verglichen.

Für die acht Repositorien wurde festgestellt, dass Videos und andere Forschungsdaten gleich behandelt werden. Für die Behandlung von Forschungsdaten im Allgemeinen konnten Gemeinsamkeiten und Unterschiede, sowie eine Vielzahl an Tendenzen identifiziert werden. Deutlich wurde, dass Gebiete auf denen Standards und Best Practice Beispiele zur Verfügung stehen (Langzeitarchivierung, Usability, Finanzierung, Barrierefreiheit, Zertifizierung), besonders große Heterogenität aufweisen. Es wurde jedoch auch festgestellt, dass internationale Standardisierungsbemühungen mindestens starke Tendenzen zur Akzeptanz zeigen. Beispiele dafür sind die Verwendung von Creative Commons Lizenzen, des DOI Systems, von Programmierschnittstellen und Dublin Core.

Abstract

The goal of this thesis was to analyse research data repositories that store videos as research data and to compare and contrast different aspects. Eight repositories were chosen for a qualitative analysis. The data was procured through an analysis of the repository websites on the basis of a re-used catalogue of criteria. The results of the analysis were then compared for each criterion.

The conclusion for the eight analysed repositories was, that videos and other types of research data are treated the same. Similarities and differences, as well as a number of tendencies, became apparent for research data regardless of type. Results show, that areas with established standards and best practices (digital preservation, usability, financing, accessibility, certification) are especially heterogeneous. It was also concluded that international standardization efforts show at least a strong tendency of being accepted. Examples are the use of Creative Commons licenses, the DOI system, Application Programming Interfaces and Dublin Core.

Inhaltsverzeichnis

1. Einleitung	1
2. Methode	5
3. Forschungsstand	11
4. Vorstellung der Forschungsdatenrepositorien	19
5. Analyse der Forschungsdatenrepositorien	22
5.1. Stanford Digital Repository	22
5.2. Phaidra	24
5.3. Zenodo	28
5.4. Brown Digital Repository	31
5.5. LSE Research Online	34
5.6. OceanRep GEOMAR Repository	37
5.7. Marine-Geo Digital Library	41
5.8. DR-NTU Data Repository	44
6. Vergleich der Forschungsdatenrepositorien	47
6.1. K1 Veröffentlichung und Nachnutzung	47
6.2. K2 Zugriff und Auffindbarkeit	50
6.3. K3 Transparenz und Qualitätssicherung	51
6.4. K4 Nutzererlebnis und Support	53
6.5. K5 Organisation und Infrastruktur	55
6.6. K6 Gebühren und Finanzierung	57
7. Bewertung der Ergebnisse	58
8. Fazit	66
8.1. Ausblick	67
9. Verzeichnisse	69
9.1. Abbildungsverzeichnis	69
9.2. Tabellenverzeichnis	69
Literatur	70
A. Anhänge	i
A.1. Analyse der Forschungsdatenrepositorien aus re3data	i
A.2. Ausgefüllte Kriterienkataloge der FDR	x
A.3. Vergleichstabellen	lxiv
A.4. Screenshots	lxx

Abkürzungsverzeichnis

BDR	Brown Digital Repository
CC	Creative Commons
CDS	Center for Digital Scholarship (Brown University)
CERN	Conseil Européen pour la Recherche Nucléaire / Europäische Organisation für Kernforschung
CITS	Centre for IT Services (Nanyang Technological University)
DC	Dublin Core
DCMI	Dublin Core Metadata Initiative
DDI	Data Documentation Initiative
DFG	Deutsche Forschungsgemeinschaft
DH	Digital Humanities
DOI	Digital Object Identifier
DRUID	Digital Resource Unique Identifier
DSA	Data Seal of Approval
EAD	Encoded Archival Description
FAQ	Frequently Asked Questions (auch: FAQs)
FD	Forschungsdaten
FDM	Forschungsdatenmanagement
FDR	Forschungsdatenrepositorium
FDZ	Forschungsdatenzentrum
FGDC	Federal Geographic Data Committee
GND	Gemeinsame Normdatei
GNU GPL	GNU General Public License (auch: GPL)
IIIF	International Image Interoperability Framework
JSON	JavaScript Object Notation
LOM	Learning Object Metadata
LSERO	LSE Research Online
LTI	Learning Tools Interoperability
LTO	Linear Tape-Open
LZA	Langzeitarchivierung
MD	Metadaten
MGDL	Marine-Geo Digital Library
MGDS	Marine Geoscience Data System
OAI-PMH	Open Archives Initiative - Protocol for Metadata Harvesting
OAIS	Open Archival Information System
ODC	Open Data Commons
OPAC	Online Public Access Catalogue

PID	Persistent Identifier (Kontext: Objekte, Bsp.: DOI)
PID	Personal Identifier (Kontext: Personen, Bsp.: ORCID)
PURL	Persistent URL
SDR	Stanford Digital Repository
URN	Uniform Resource Name
ZID	Zentraler Informatikdienst (Universität Wien)

1 Einleitung

"Knowledge is open if anyone is free to access, use, modify, and share it — subject, at most, to measures that preserve provenance and openness." [Ope15]

Die Offen-Definition ist eine der Grundlagen der Vielzahl an offenen Bewegungen, die derzeit die Arbeitsweisen von Wissenschaftler¹ grundlegend verändern. Open Science, Open Access, Open Data und Open Source sind die größten Strömungen dieser neuen Philosophie des offenen Austausches. Die darin vertretenen Ansätze werden zunehmend von Regierungsorganisationen (Open Research Data Pilot der EU², OECD³, White House Office of Science and Technology Policy⁴, Bundesministeriums für Bildung und Forschung⁵), Wissenschaftsorganisationen (Deutsche Forschungsgemeinschaft⁶, Allianz der deutschen Wissenschaftsorganisationen⁷, US National Science Foundation⁸, Rat Für Informationsinfrastrukturen⁹), Forschungsförderern (Europäische Kommission¹⁰, Überblick des Digital Curation Center¹¹) und Zeitschriftenherausgebern (Public Library of Science¹², SpringerNature¹³) unterstützt. Einige dieser Policies verpflichten zur offenen Publikation von Forschungsdaten.

Im Zusammenhang mit Open Science und Open Data wird immer wieder die Frage laut, wo genau Forschungsdaten abgelegt werden können. Derzeit werden Forschungsdatenrepositorien als Antwort gehandelt; sie haben das Ziel die Daten dauerhaft auffindbar und zugänglich zu machen. Gleichzeitig wird es für Forscher^x immer schwieriger einen Überblick über die große Anzahl von Forschungsdatenrepositorien zu gewinnen.

Motivation

In dieser Arbeit sollen Forschungsdatenrepositorien untersucht werden, die audiovisuelle Forschungsdaten beinhalten. Eine der Feststellungen von Eidenberger fasst die Motivation zur

¹In dieser Arbeit werden Angehörige aller Gender angesprochen und nach der x-Form aus AG Feministisch Sprachhandeln (Hg.) (2015): Was Tun? Sprachhandeln - aber wie? W_Ortungen statt Tatenlosigkeit. 2. Aufl. Humboldt-Universität zu Berlin. Berlin. Online verfügbar unter http://feministisch-sprachhandeln.org/wp-content/uploads/2015/04/sprachleitfaden_zweite_auflage.pdf gegendert (ab Seite 22). Beispiel: "Forscher^x" wird als "Forscher^{ix}" gelesen / gesprochen.

²<https://data.europa.eu/euodp/de/data/dataset/open-research-data-the-uptake-of-the%2Dpilot-in-the-first-calls-of-horizon-2020>

³<https://www.oecd.org/sti/sci-tech/38500813.pdf>

⁴<https://obamawhitehouse.archives.gov/blog/2016/02/22/increasing-access-results%2Dfederally-funded-science>

⁵<https://www.bmbf.de/foerderungen/bekanntmachung.php?B=774>

⁶http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf

⁷<https://www.allianzinitiative.de/de/archiv/forschungsdaten/grundsaeetze.html>

⁸<https://www.nsf.gov/bfa/dias/policy/dmp.jsp>

⁹<http://www.rfii.de/de/category/dokumente/>

¹⁰https://ec.europa.eu/research/science-society/document_library/pdf_06/recommendation-access-and-preservation-scientific-information_en.pdf

¹¹<http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>

¹²<http://journals.plos.org/plosone/s/data-availability>

¹³<http://www.springernature.com/de/authors/research-data-policy/data-policy-types/12327096/>

Beschäftigung mit audiovisuellen Medien gut zusammen: "The visual sense is of particular interest for multimedia information retrieval. The digital culture of the 21st century is a visual one." [Eid12, S. 570]. Dies folgt der Beobachtung, die auch im Horizon Report dargestellt wurde, nämlich dass ein medialer, gesellschaftlicher Trend in Richtung Video, Visualisierungen und Virtual Reality besteht [Ada+17, S. 2]. Auch die nestor-AG Media stellt die Entwicklung fest: "In allen Kultur- und Bildungsbereichen entstehen zunehmend digitale Materialien und Informationen, die nicht nur textbasiert sind, sondern Elemente wie Ton, Grafik, Video, Animation, Simulation etc. enthalten [...]" [nK17]. Eidenberger fügt in seinem Handbuch zum Thema Multimedia Retrieval dazu: "Information retrieval experts are mostly concerned with the processing and understanding of text." [Eid12, S. 13], welches sich auch auf die Untersuchungen zum Umgang mit audiovisuellen Forschungsdaten in FDR auszuweiten scheint.

Als Besonderheiten der Behandlung digitaler audiovisueller Daten nennt Deggeller die schnelllebigen Hard- und Softwareumgebungen, große Datenmengen, verlustreiche und den Bitstream beschädigende Algorithmen zur Reduzierung eben dieser Datenmengen und die kurze Lebensdauer der Formate durch fehlende internationale Standards. Er fasst zusammen: "Es übersteigt die Möglichkeiten der meisten Archive, Bibliotheken und Museen, den Überblick über diese Entwicklung zu behalten." [Deg14, S. 48]. Die nestor-AG Media pflichtet Deggeller vor allem beim Größenproblem zu: "Ein digitaler Film, der ohne Qualitätsverlust gespeichert wird, benötigt schnell mehrere Terabyte an Speicherplatz." [nes16, S. 6]. Als Hauptproblem werden ebenfalls die proprietären und kurzlebigen Formate identifiziert. Dabei beziehen sich beide Publikationen auf audiovisuelle Medien im Allgemeinen und nicht auf FD im Besonderen, dennoch ist der Weg zum Umgang mit Forschungsdaten wegen der gemeinsamen Problemfelder ein kurzer. Die Arbeit von Reinhold, Rittberger und Keil, die sich mit audiovisuellen Forschungsdaten beschäftigt, pflichtet dem Größen- und Übersichtsproblem zu, nennt aber auch den Schutz sensibler Daten und den erhöhten Produktionsaufwand als besondere Herausforderungen von wissenschaftlichen Videos [RRK13, S. 201].

Ziel der Arbeit

Der Umgang mit audiovisuellen Forschungsdaten in Repositorien ist wegen den Besonderheiten der Ressource interessant. Ob Forschungsdatenrepositorien diese entsprechend gesondert behandeln, soll für einige ausgewählte Fälle überprüft werden. Dabei sind Videodaten von vorrangigem Interesse, werden aber gemeinsam mit allen anderen Ressourcen im Repitorium betrachtet.

Die Fragen, die am Ende dieser Arbeit beantwortet werden sollen, sind:

- Wie gehen die ausgewählten Repositorien mit audiovisuellen Forschungsdaten um? Wie gehen sie mit Forschungsdaten im Allgemeinen um?
- Gibt es Gemeinsamkeiten und Unterschiede in der Behandlung von Videos oder Forschungsdaten generell?
- Ist die Behandlung von Videos und anderen Forschungsdaten auf den Webseiten der Forschungsdatenrepositorien transparent dargestellt?
- Können Nutzerx anhand der Informationen auf den Webseiten entscheiden ob sie ihre Videos dort ablegen?

Begriffsklärung

Die zentralen Begriffe dieser Arbeit lassen Spielraum für Interpretation und werden zur besseren Verständlichkeit kurz definiert.

Deggeller definitert **audiovisuelle Daten** als:

"[...] Dokumente, deren Inhalt mit Tönen und/oder bewegten Bildern dargestellt ist: Tonaufnahmen, stumme oder vertonte Filme und Videos."

[Deg14, keine Seitenzahl im ebook]

Für **Forschungsdaten** wird die Definition von Bauer et al genutzt:

"[U]nter dem Begriff 'Forschungsdaten' [sind] alle Daten zu verstehen, die im Zuge wissenschaftlicher Forschungs- und künstlerischer Schaffensprozesse entstehen (z.B. Text, Tabellen, Video, Audio, Grafik etc.) und auf deren Grundlage ihre Forschungsergebnisse und/oder Kunstwerke basieren – z.B. durch Experimente, Quellenforschungen, Messungen, Erhebungen, Digitalisate oder Entwürfe."

[Bau+15, S. 15]

Die nestor-AG Medien stellt fest: "Prinzipiell kann jedes audiovisuelle Werk wissenschaftlich genutzt werden." [nes16, S. 43]

Im Allgemeinen folgt diese Arbeit der Definition für **Repositorien** von Becker:

"Ein Repositorium ist ein System zur sicheren Speicherung und Weitergabe digitaler Objekte und der die Objekte beschreibenden strukturierten Metadaten."

[Bec14, S. 16]

Beziehungweise der Definition von **Forschungsdatenrepositorien** von Kindling et al:

"A [research data repository] is a technical and organizational information system that helps researchers to manage, store and provide their own datasets, and to easily find and access datasets from other sources."

[Kin+17, S. 2]

Thematische Eingrenzung

In dieser Arbeit kann nicht auf die Besonderheiten von allen Arten von audiovisuellen Forschungsdaten eingegangen werden. Die Betrachtung aller audiovisuellen Ressourcen, also auch Ton oder reines Bewegtbild, ist zu umfassend. Ein Beispiel: Bewegtbilder wie GIFs unterscheiden sich in ihren Ansprüchen und Möglichkeiten stark von Videos. Sie liegen in anderen Formaten vor, sind nicht vergleichbar im Datenvolumen und haben andere Herausforderungen in der Archivierung. Daher werden in dieser Arbeit nur solche Ressourcen betrachtet, die ein bewegtes Bild mit einer Tonspur verbinden. Dies schließt Ressourcen aus, die nur aus bewegten Bildern bestehen (z.B. Dateityp .gif) oder nur Tonaufnahmen sind (z.B. .mp3, .wav). Verkürzt wird für die betrachteten Ressourcen in dieser Arbeit auch der Begriff "Video" verwendet. Primär sind Online verfügbare Videos von Interesse; ob es sich dabei um Digitalisate in Dateiform oder born digital Videos handelt wird in diesem Kontext nicht untersucht.

Alle Repositorien mit audiovisuellen Forschungsdaten zu untersuchen, würde den Rahmen der vorliegenden Arbeit überschreiten. Daher wurde eine Anzahl von FDR ausgewählt um daran

eine beispielhafte Analyse vorzunehmen. Mehr zur Auswahl der Repositorien in Kapitel 2. Die gewählten Repositorien sind nicht repräsentativ und es wird kein Anspruch auf einen umfassenden Überblick gelegt.

Aus Praktikabilitätsgründen kann diese Arbeit keine Betrachtung aller Stakeholder des FDM und derer Ansprüche an Repositorien leisten. Daher wird die Perspektive des Nutzerx eingenommen, dabei sind datengebende Personen genauso inbegriffen wie Datennachnutzende.

Es ist zu beachten, dass diese Arbeit auf einem Verständnis von Forschungsdatenmanagement im bibliothekarischen Sinne basiert. Die theoretischen Grundlagen werden nicht noch einmal erläutert. Stattdessen sei an dieser Stelle auf [BHM11] (darin für FDR [AN11]) und [Neu+12] verwiesen. Für eine Data Librarian Perspektive wird [RS16] empfohlen. Eine Einführung in die Funktionen von Forschungsdatenrepositorien¹⁴ ist auf [Forschungsdaten.info](https://www.forschungsdaten.info) verfügbar.

In dieser Arbeit wird das tatsächliche Retrieval von audiovisuellen Daten bzw. deren technische Anforderungen nicht behandelt, eine sehr umfassende Publikation zu diesem Thema verfasste [Eid12].

Aufbau der Arbeit

In Kapitel 1 wurde beschrieben, warum Repositorien mit audiovisuellen Forschungsdaten untersucht werden und wie die zentralen Begriffe der Arbeit verwendet werden. In Kapitel 2 wird erläutert wie die Recherche durchgeführt wurde und wie die zu untersuchenden Repositorien ausgesucht wurden. Die Auswahl des nachgenutzten Kriterienkataloges wird erläutert. Im Forschungsstand in Kapitel 3 werden einige Repositorienverzeichnisse verglichen und Arbeiten betrachtet, die sich ebenfalls mit der Analyse und dem Vergleich von Repositorien und Forschungsdatenrepositorien beschäftigten. Außerdem wird eine Auswahl relevanter Aspekte bei der Behandlung von Videos vorgestellt. Kapitel 4 enthält eine kurze Vorstellung aller analysierten Repositorien als Einführung in die Ergebnisse dieser Arbeit. Zuerst wird in Kapitel 5 für jedes FDR eine Zusammenfassung der Analyse der Webseiten präsentiert. Darauf aufbauend erfolgt in Kapitel 6 ein Vergleich der Ergebnisse der Analyse für jedes der Kriterien. Die Bewertung der Ergebnisse in Kapitel 7 beinhaltet eine Zusammenfassung aller Ergebnisse, sowie eine Betrachtung der Kriterien, die für audiovisuelle Forschungsdaten als besonders identifiziert wurden. Die Arbeit schließt in Kapitel 8 mit einem Fazit und gibt einen Ausblick auf mögliche Änderungen in der Behandlung von audiovisuellen Forschungsdaten.

Die Anhänge enthalten vor allem die Forschungsdaten dieser Arbeit. Anhang A.1 enthält eine Liste von FDR aus der ersten Analyse aus re3data und daraus resultierende Ergebnisse. Darauf folgen in Anhang A.2 die ausgefüllten Kriterienkataloge für alle FDR in nicht zusammengefasster Form. Die Tabellen in Anhang A.3 waren eine Arbeitshilfe zur Erstellung des Vergleichskapitels (Kap. 6). Nominell als Anhang A.4 bezeichnet liegen die Screenshots der Webseiten aus Platzgründen auf dem Datenträger bei.

¹⁴<https://www.forschungsdaten.info/themen/bewahren-und-nachnutzen/repositorien/>

2 Methode

Für die Suche nach Literatur wurde die Methode des Literature Review angewandt. Dazu wurden aktuelle Jahrgänge diverser Fachzeitschriften und die Informationsquellen PubLIS, BASE, INFODATA-eDepot, LISA, e-LIS, Google Scholar, VÖBB und BIB Info OPUS nach verschiedenen Aspekten von Forschungsdatenrepositorien (Vergleich von, Beschreibung von, Kriterien für) und möglichen Besonderheiten von audiovisuellen Forschungsdaten durchsucht.

In der Recherche wurde auf Publikationen der letzten 5 Jahre eingeschränkt, da die technische Entwicklung auf diesem Gebiet die Projekte und Publikationen zu Repositorien schnell veralten lässt. Ausgenommen davon sind Grundlagenwerke des FDM und Primärquellen für die Analyse der FDR. Hochschulschriften (inkl. online zugänglichen Bachelorarbeiten), graue Literatur, online verfügbare Konferenzfolien und Webquellen wurden explizit mit in die Recherche einbezogen.

Die Filterung der Ergebnisse geschah mehrstufig: In der ersten Stufe erfolgte eine reine Titelanalyse. Die daraus extrahierten Arbeiten wurden anhand des Abstracts bewertet. Darauf folgte eine inhaltliche Analyse der übrig gebliebenen Arbeiten, und daraus schließlich eine Extraktion der relevanten Angaben. Zusätzlich wurde für Arbeiten der letzten Stufe das Literaturverzeichnis anhand der Schneeball-Technik nach weiterer relevanter Literatur durchsucht.

Für die Untersuchung der Repositorien wurde die Methode der "Desktop-Recherche" [VW14] bzw. der qualitativen Inhaltsanalyse von Webseiten [VK14] angewandt, d.h. die Web-Auftritte der FDR dienten als primäre Informationsquelle. Online verfügbare Publikationen wurden nur unterstützend verwendet. Zu beachten ist bei diesem Vorgehen, dass das Fehlen von Informationen zu einem Kriterium auf der Webseite keine Aussage zum tatsächlichen Angebot des Repositoriums trifft. Diese Methode bietet sich dennoch an, da in der vorliegenden Arbeit die Sicht der Forschenden eingenommen wird. Diese nutzen ebenfalls die Webauftritte der FDR als primäre Informationsquelle, um ein FDR für die Aufbewahrung der eigenen Forschungsdaten auszuwählen.

Auswahl der Forschungsdatenrepositorien

Im Laufe dieser Arbeit wurde eine Auswahl von Forschungsdatenrepositorien untersucht, die sich mit audiovisuellen FD auseinandersetzen. Hierbei wurde kein Anspruch auf Vollständigkeit oder Repräsentativität gelegt, stattdessen wurden die FDR qualitativ untersucht. Die Auswahl geschah anhand des Repositorienverzeichnisses re3data. Mehr Informationen zu Repositorienverzeichnissen in Kapitel 3.

Entsprechend dem Fokus auf audiovisuelle FD wurde an erster Stelle für die Facette contentTypes "audiovisual data" festgelegt. Die Behandlung weiterer Datentypen in den gewählten FDR ist dabei zu erwarten und stellte kein Ausschlusskriterium dar. Die Abbildung 2.1 visualisiert die Auswahl der FDR.

In der Abbildung steht das Rechteck für den Grundzustand in re3data, Rauten visualisieren die getroffenen Entscheidungen anhand der angewandten Filter und Parallelogramme beinhalten den Output der Entscheidungen, also die Anzahl der verbleibenden FDR. Die Pfeile weisen darauf hin, dass der Prozess iterativ stattfand. Die unterbrochenen Pfeile verbinden einzelne Schritte mit zusätzlichen Informationen aus dem Prozess. Das erste Zwischenergebnis und der letzte Schritt der Auswahl sind als Listen visualisiert.

VERGLEICH UND ANALYSE VON FORSCHUNGSDATENREPOSITORIEN

2. Methode

Re3data.org führte am Stichtag (22. November 2017) 1.987 FDR und darunter 375 Repositorien die audiovisuelle Daten verzeichneten. Um dies für eine qualitative Untersuchung weiter einzuschränken wurden zusätzliche Kriterien aus dem re3data Schema [Rüc+15] bestimmt.

Die Facette "qualityManagement" wurde gewählt um sicherzustellen, dass nur FDR untersucht werden, deren FD einen durch Qualitätsmanagement gesicherten hohen Standard an Forschungs- oder Metadaten haben. Rice & Southall stellten fest, dass jede Art von Qualitätssicherung vor dem Einspielen der Daten später "time, effort and frustration" [RS16] (S. 113) spart.

Drei Facetten wurden festgelegt um sicherzustellen, dass Zugang zum FDR und zu den enthaltenen FD besteht: Für "databaseAccessType" wurde die Angabe Open gewählt um FDR auszuschließen, die als Zugang zur Datenbank Closed und Restricted angeben. Jedes FDR kann nur eine Art von Zugang angeben [Rüc+15, S. 10].

Für "dataAccessType" wurde ebenfalls Open gewählt um sicherzustellen, dass die FDR Zugang zu mindestens einem Datensatz erlauben. Da jedoch Mehrfachangaben in diesem Feld erlaubt sind, können die FDR auch Closed, Embargoed und Restricted Datensätze enthalten. Die dritte Facette "dataAccessRestrictions" wurde auf Registration eingeschränkt um damit

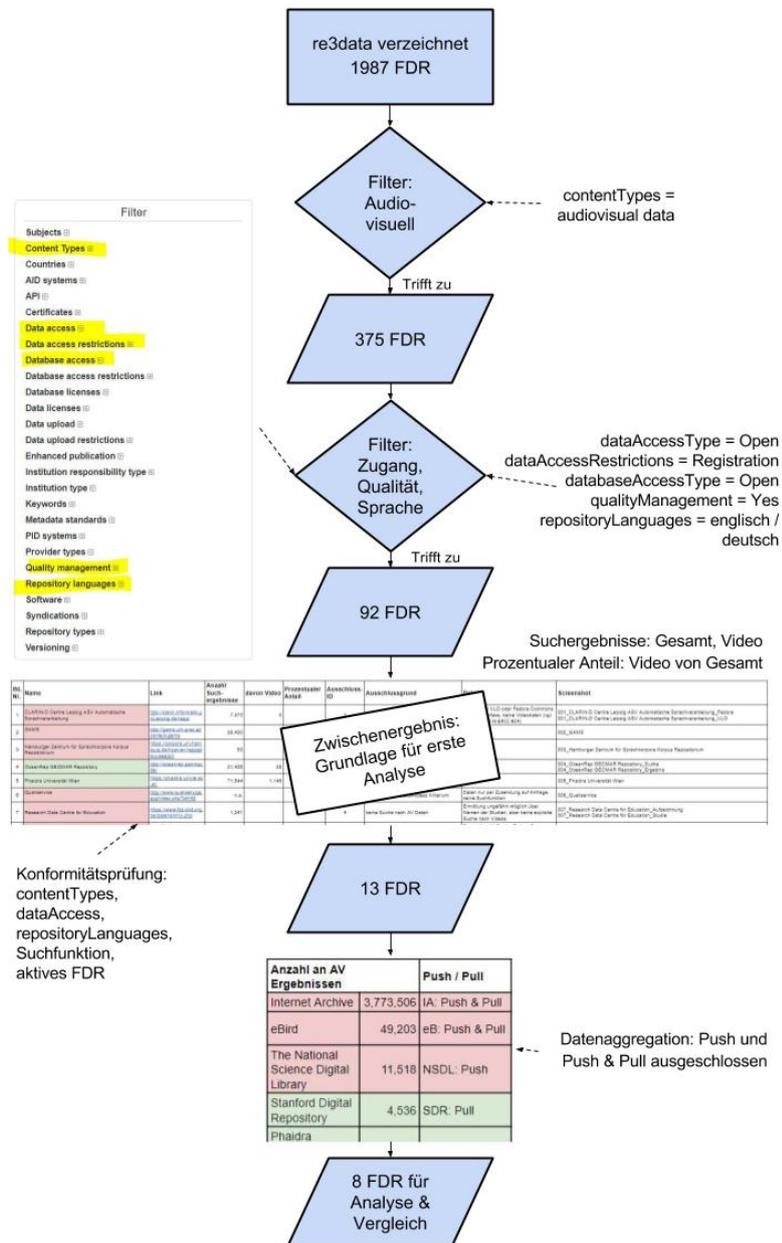
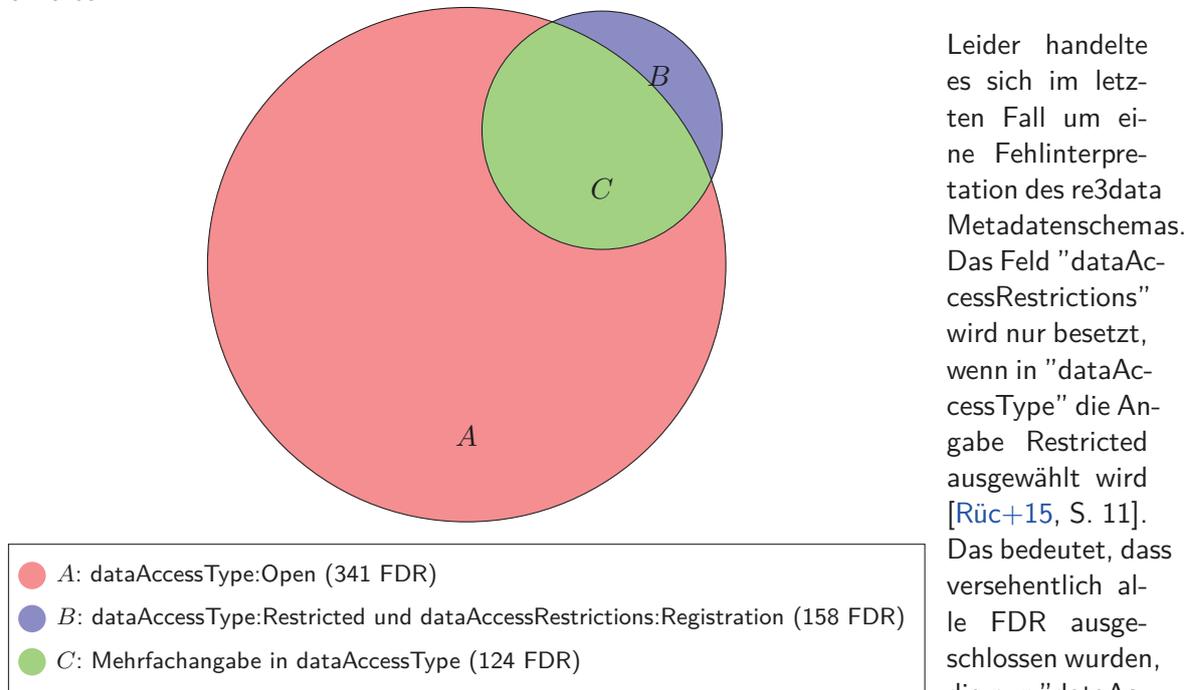


Abbildung 2.1.: Auswahl der zu untersuchenden FDR (eigene Darstellung).

FDR auszuschließen, die eine Bezahlschranke haben oder eine institutionelle Zugehörigkeit erwarten.



Leider handelte es sich im letzten Fall um eine Fehlinterpretation des re3data Metadatenschemas. Das Feld "dataAccessRestrictions" wird nur besetzt, wenn in "dataAccessType" die Angabe Restricted ausgewählt wird [Rüc+15, S. 11]. Das bedeutet, dass versehentlich alle FDR ausgeschlossen wurden, die nur "dataAccessType:Open" angegeben haben (und dadurch das Feld "dataAccessRestrictions" leer bleibt), in der

Abbildung 2.2.: Anzahl von FDR nach dataAccessType und dataAccessRestrictions in re3data, Stand: Januar 2018¹⁵(eigene Darstellung).

Abbildung 2.2 mit "A" beschriftet. Beispiele für FDR die dadurch ausgeschlossen wurden sind unter anderem "The Cell Image Library" und "PANGAEA". Dadurch, dass "dataAccessType:Open" zuerst gewählt wurde, beinhalteten die Ergebnisse alle Repositorien die zusätzlich zu Open noch weitere Arten von Zugriffsbeschränkungen anbieten, darunter auch die Menge "C" in der die FDR "dataAccess:Restricted" und "dataAccessRestrictions:registration" melden.

Durch die Möglichkeit der Mehrfachangabe in "dataAccessType" handelt es sich bei der hier getroffenen Auswahl (in der Abbildung mit "C" markiert) also um eine Untermenge der ursprünglich zur Untersuchung vorgesehenen Repositorien (in der Abbildung mit "A" markiert), die ebenfalls "C" beinhaltet hätte. Die Menge "B" enthält FDR die nur "dataAccess:Restricted" mit der Einschränkung "dataAccessType:Registration" angeben, aber keinen offenen Datenzugang.

Einige Ergebnisse der Analyse und des Vergleichs wurden durch die Auswahl der FDR beeinflusst, dies wird in der Interpretation der Ergebnisse jedoch berücksichtigt. Der Fehler wurde erst nach Abschluss der inhaltlichen Analyse der ausgewählten FDR entdeckt. Aus Zeitgründen war eine Neuerhebung auf Grundlage der ursprünglich geplanten Menge "A" nicht mehr möglich.

Als weiteres Kriterium wurde "repositoryLanguage" mit "deu" oder "eng" (deutsch- oder eng-

¹⁵Die in der Abbildung enthaltenen Zahlen entsprechen dem Stand Januar 2018 und weichen von den Zahlen am Stichtag im November 2017 ab, da diese nicht mehr rekonstruiert werden konnten. Sie dienen aber nur der Visualisierung der Schnittmengen und deren Relation zueinander.

lischsprachiges Userinterface) festgelegt. Die Einschränkung auf die Sprachen wurde vorgenommen, um die Wahrscheinlichkeit einer für die Autorin lesbaren Dokumentation und Webseite zu erhöhen.

Diese Auswahl ergab 10 als deutschsprachig und 90 als englischsprachig markierte Repositorien (davon 8 mit beiden Sprachen markiert). Im nächsten Schritt wurden alle 92 FDR hinsichtlich der Einhaltung der Kriterien und dem prozentualen Anteil der audiovisuellen FD analysiert (in Abbildung 2.1 als "Zwischenergebnis" markiert).

In dieser ersten Analyse wurde die fehlende Möglichkeit der Einschränkung auf audiovisuelle Daten sowie das Fehlen eines adäquaten Suchformulars oder fehlender Zugang zur Suche als weitere Ausschlusskriterien gesetzt. Die Verteilung von Ausschlussgründen die angewendet wurden, ist in Tabelle 2.1 aufgeführt.

Eine volle Liste der FDR inkl. der Ausschlussgründe und weiteren relevanten Details befindet sich im Anhang A.1. Screenshots der Webseiten, die die in der Tabelle genannte Fakten belegen, liegen auf dem Datenträger im Ordner "Analyse der 92 FDR aus re3data" bei.

Von den 92 FDR erfüllten nur 13 Repositorien (also knapp 12%) die oben festgelegten Kriterien. Dies beinhaltet eine Abweichung von den in re3data hinterlegten Angaben (contentType, dataType, repositoryLanguages) in 15% der Fälle (17 von 92).

Im letzten Schritt wurde die Art der Datenkollektion geprüft, um einen Vergleich zwischen zwei grundlegend verschiedenen Arten von Repositorien zu vermeiden. In dieser Arbeit werden Push-aggregierende FDR untersucht, da die Perspektive der Nutzerx eingenommen wird, die dort ihre Daten selbst hochladen können. Aus der vorigen Ergebnisliste (siehe Tabelle A.4 in Anhang A.1) wurden daher 5 Repositorien ausgeschlossen, die ihre Daten in einem automatischen oder Push Verfahren aggregieren (z.B. durch Crawler) statt durch manuelle Eingabe bzw. Pull Verfahren (z.B. direkter Upload durch Nutzerx). Ausgeschlossen wurden auch solche FDR die Push und Pull Verfahren nutzen. Dies war über die Kriterien in re3data nicht möglich, da Push / Pull nicht zwingend mit den verzeichneten OAIS-Konzepten von DataProvider und ServiceProvider einhergeht. Die Screenshots aller 13 Webseiten, die die Datenkollektion belegen, liegen auf dem Datenträger im Ordner "Push-Pull Prüfung" bei. Ausgeschlossen wurde die National Science Digital Library da sie nur Pull-Verfahren nutzt, sowie das Internet Archive, eBird, die UCD Digital Library und das UK Data Archive die alle Push- und Pull Verfahren nutzen um Daten zu aggregieren.

Nach Eingrenzung auf Repositorien in denen FD durch Nutzerx oder Mitarbeiterx eingespeist

Ausschlussgrund	Anzahl
Entspricht nicht contentType Kriterium	9
Entspricht nicht dataType Kriterium	2
Entspricht nicht repositoryLanguages Kriterium	6
keine Suche nach AV Daten	39
keine passende Suchfunktion	10
FDR Offline	2
kein FDR auf Webseite	5
nur Portal, Datenlieferant separat betrachtet	6

Tabelle 2.1.: Ausschlussgründe und deren Anzahl in der ersten Analyse der Repositorien

werden, verblieben 8 FDR für die tiefere Analyse und den Vergleich. Die Liste der untersuchten FDR umfasst:

- Stanford Digital Repository
- Phaidra
- Zenodo
- Brown Digital Repository
- London School of Economics and Political Science Research Online
- OceanRep GEOMAR Repository
- Marine-Geo Digital Library
- DR-NTU Data repository

Vergleichskriterien

Um die FDR auf einer angemessenen Ebene zu vergleichen wurde die Analyse der Webseiten anhand eines Kriterienkataloges vorgenommen. Auf eine neue Erarbeitung eines solchen Kataloges wurde zugunsten der Nachnutzung bereits bestehender Kataloge verzichtet.

Eine mögliche Grundlage waren die Checklisten, anhand derer Repositorien zertifiziert werden. Das "Memorandum of Understanding to create a European Framework for Audit and Certification of Digital Repositories" [GHK10] verständigte sich auf die drei Ebenen der Zertifizierung von Repositorien: Basic, Extended und Formal Certification. Dabei handelt es sich bei der Basic Certification um die Erlangung des Data Seal of Approval [DAN10]¹⁶. Für eine Extended Certification muss das DSA erlangt werden und zusätzlich eine öffentlich verfügbare Selbstauskunft auf Basis von ISO 16363¹⁷ oder DIN 31644¹⁸ durchgeführt werden. Das höchste Zertifikat "formal" verlangt neben dem DSA eine externe Prüfung nach ISO 16363 oder DIN 31644 [GHK10].

Ebenfalls erwägt wurden der "Kriterienkatalog vertrauenswürdige digitale Langzeitarchive" der nector Arbeitsgruppe Vertrauenswürdige Archive¹⁹, die "DRAMBORA - Digital Repository Audit Method Based on Risk Assessment" des Digital Curation Centre²⁰, die "Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC)" des Center for Research Libraries²¹ und die "Audit and certification of trustworthy digital repositories" des Consultative Committee for Space Data Systems (ISO16363:2012)²². Ein Großteil davon ist jedoch für die Selbstevaluierung von Repositorien konzipiert: Die Zertifizierungen nehmen eher die Sicht der Betreiberx als die der Nutzerx ein und erfragen zum Teil Informationen die nur von Mitarbeiterx beantwortet werden können.

Um die FDR zu analysieren werden stattdessen die von Steinhof entwickelten "Erfolgskriterien für Forschungsdatenrepositorien" [Ste17] nachgenutzt. Die dort erarbeiteten Kriterien basieren auf vier Dokumenten: "Policy-making for Research Data in Repositories" aus dem DISC-UK

¹⁶<https://www.datasealofapproval.org/en/information/requirements/>

¹⁷Auch: "Audit and certification of trustworthy digital repositories", <https://www.iso.org/standard/56510.html>

¹⁸Auch: "nector-Siegel für vertrauenswürdige digitale Langzeitarchive", <https://www.din.de/de/mitwirken/normenausschuesse/nid/normen/wdc-beuth:din21:147058907?sourceLanguage&destinationLanguage>

¹⁹<https://edoc.hu-berlin.de/handle/18452/2175>

²⁰<http://www.repositoryaudit.eu/about/>

²¹<http://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and%2Dcertifying/trac>

²²<https://public.ccsds.org/pubs/652x0m1.pdf>

Projekt [GMR09], der "Matrix of use cases and functional requirements for research data repository platforms" der Research Data Alliance [RDA16], dem Artikel "Are Scientific Data Repositories Coping with Research Data Publishing?" [Ass+16] und dem re3data.org "Schema for the description of research data repositories" in der Version 2.1 [Vie+14]. Die Arbeit betrachtet einzelne Stakeholder des FDM getrennt, darunter auch die Nutzerx von FDR.

Der Katalog ist in die Gruppen "Veröffentlichung und Nachnutzung", "Zugriff und Auffindbarkeit", "Transparenz und Qualitätssicherung", "Nutzererlebnis und Support", "Organisation und Infrastruktur" und "Gebühren und Finanzierung" unterteilt, in denen in weiteren Unterpunkten verschiedene Aspekte eines FDR thematisiert werden. Aus dem erklärenden Fließtext der Unterpunkte von Steinhof wurden weitere Aspekte identifiziert. Der in dieser Arbeit final verwendete Kriterienkatalog ist also in drei statt nur zwei Ebenen unterteilt. Eine Übersicht ist in Anhang A.2 enthalten.

Zur automatischen Untersuchung der Webseiten auf Barrierefreiheit (K4.1 Usability) wird die Browsererweiterung aXe²³ im Browser Chrome benutzt. Dabei ist zu beachten, dass die Browsererweiterung anhand der aktuellen Anzeige untersucht, d.h. für jeden aufgeklappten Tabellenreiter (z.B. bei der Anzeige von Metadaten) erscheinen die jeweiligen Verstöße. Da die Barrierefreiheit nur einen kleinen Teilaspekt der Arbeit darstellt wurden pro Repository nur zwei Unterseiten analysiert.

²³<https://www.deque.com/products/axe/>

3 Forschungsstand

Die zentralen Themen der Arbeit werden im Folgenden getrennt betrachtet, da in der Recherche wenige Überschneidungen gefunden wurden. Zuerst werden Arbeiten und Projekte auf den allgemeineren Teilgebieten betrachtet, daraufhin die Spezifischeren.

Repositorienverzeichnisse

Für Forscherx, die ein FDR für ihre Daten suchen, wurden diverse Verzeichnisse erstellt, die einen Überblick über die Vielzahl an Repositorien ermöglichen sollen. Darunter zum Beispiel OpenDOAR²⁴, OpenAire²⁵ und re3data²⁶. OpenDOAR verzeichnete Ende 2017 über 3000 Repositorien, erlaubt aber keine Einschränkung auf solche die audiovisuelle Forschungsdaten beinhalten. OpenAIRE erlaubt zwar eine typenspezifische Suche und verzeichnet über 4000 passende FDR, erschwert daraufhin aber eine weitere Eingrenzung über die eingeschränkte Facettierung. Da re3data beides bietet wurde es als Grundlage für die Auswahl der FDR bestimmt. Es ist weiterhin geeignet da es sich um ein internationales Register handelt, dass nur FDR listet, die ihre Zugangs- und Nutzungsvoraussetzungen offenlegen und von einer Institution getragen werden, die den langfristigen Zugang zu den Forschungsdaten sichern können [Rüc+15]. Das zugrundeliegende Metadatenschema wird auf der re3data Webseite zu großen Teilen als Facettierung für die Suche zur Verfügung gestellt und erlaubt somit eine sehr konkrete Auswahl. Weitere Details zu re3data bietet z.B. Forschungsdaten.org²⁷. Der Vollständigkeit halber soll die Liste von DINI-zertifizierten Repositorien²⁸ erwähnt werden, die für Repositorien im deutschen Raum zentral ist. Aufgrund der fehlenden Einschränkungsmöglichkeiten wurde die DINI-Liste für diese Arbeit nicht verwendet.

Vergleich von Repositorien anhand eines Kriterienkatalogs

In der Masterarbeit von Offhaus (2012) wurden dreizehn institutionelle Repositorien von Nordrhein-Westfälischen Universitäten anhand eines Kriterienkataloges untersucht. Dieser umfasst vierzehn Kriterien unterteilt in drei Anforderungsgruppen: Allgemein, datentechnisch und bibliothekarisch [Off12, S. 30]. Die Kriterien basieren auf einer Analyse von Fachliteratur, darunter primär den Empfehlungen des Wissenschaftsrates abgeglichen mit dem DINI-Zertifikat in der Fassung von 2010. Daran erfolgte eine SOLL/IST-Analyse (Anforderungen / tatsächlich umgesetzte Angebote), die eine Übereinstimmung von fast 42% für die Allgemeinen, sowie 64% für die Datentechnischen und 27% für die bibliothekarischen Anforderungen ergab. Insgesamt wurde von Offhaus eine Umsetzung von etwas über 40% der erarbeiteten Anforderungen festgestellt, mit besonders niedrigen Übereinstimmungen auf den Gebieten "Etablierung, Pflege und Betrieb institutioneller Repositorien" (ebd., S.46). Es wurde u.a. empfohlen nutzerfreundliche Leitlinien anhand des DINI-Zertifikates zu formulieren (ebd., S. 50) und zum Vertrauensgewinn

²⁴<http://www.opendoar.org/find.php>

²⁵<https://www.openaire.eu/>

²⁶<https://www.re3data.org/>

²⁷<http://www.forschungsdaten.org/index.php/Re3data.org>

²⁸<https://dini.de/dini-zertifikat/liste-der-repositorien/>

eine Zertifizierung anzustreben (ebd., S. 51).

Ein Vergleich einiger Open Source Softwareplattformen für Repositorien wurde 2014 von Pырounakis, Nikolaidou und Hatzopoulos für DSpace, Fedora, Greenstone, EPrints und Invenio durchgeführt. Deren Eignung wurde anhand von fünf Case Studies typischen Praxis-Situationen zugeordnet. Zu den Case Studies zählten auch die hier betrachteten "Scientific data collection" und "New media art collection" [PNH14, S. 15] in denen Bilder, Video und Audio Dateien erwartet werden. Die Autoren stellten auf Basis des "Reference Model for an Open Archival Information System"²⁹ (OAIS) und dem "DELOS DL Reference Model"³⁰ vierzehn Kriterien für den Vergleich auf. Darunter Umgang mit Metadaten, Suchfunktion, Zugangskontrolle und Schnittstellen (ebd., S. 3f). Für FD empfehlen die Autoren Invenio und DSpace (ebd., S. 14). Für multimediale Objekte, wie die hier betrachteten Videos, werden DSpace und Fedora empfohlen (ebd., S. 15f). Sie kamen zu dem Schluß, dass alle fünf Softwareplattformen die nötigen "Kernfunktionen" (ebd., S. 16) aufwiesen, jedoch keine alle aufgestellten Kriterien abdeckte. Voigt untersuchte 15 deutsche, institutionelle Open Access Repositorien mit einem Fokus auf Rechteübertragung und Lizenzmanagement. Die Auswahl der Repositorien geschah auf Grundlage von BASE. Die zu untersuchenden Repositorien wurden über Kriterien und Schwellenwerte (z.B. Volltextzugriff für >50% der Publikationen) identifiziert [Voi14, S. 29]. Aus dieser Auswahl wurden per Zufallsgenerator zehn institutionelle und von DINI zertifizierte und fünf weitere Repositorien ermittelt. Die Autorin begann mit einer Inhaltsanalyse der Webseiten auf Basis einer Checkliste mit Kriterien (ebd., S. 30). Die dort nicht vorhandenen Informationen wurden per E-Mail von den Repositorienbetreibern erfragt (ebd., S. 31). Zusätzlich wurden eine Auswahl von relevanten Metadaten über die jeweiligen OAI-PMH Schnittstellen ausgelesen und analysiert (ebd., S. 31f). Dies ergab u.a. dass Creative Commons Lizenzen etabliert sind (ebd., S. 56), fast nur einfache Nutzungsrechte übertragen werden (ebd., S. 55) und die Metadatenqualität der Lizenzangaben sehr heterogen ist (ebd. S. 56). Aus den Ergebnissen wurden acht Handlungsempfehlungen formuliert. Empfohlen werden u.a. die Verwendung von offenen Lizenzen und die Angabe von Lizenzinformationen in Metadaten und bereitgestellten Volltexten (ebd., S. 60ff).

Die Bachelorarbeit von Müller (2015) untersuchte Repositorien mit Fokus auf das Dateiformat PDF und verfolgte dabei einen quantitativen Ansatz. Ziel der Untersuchung war eine Übersicht der PDF-Versionen die in der Praxis genutzt werden. Der Autor fragte über die jeweilige OAI-PMH-Schnittstelle alle PDF-Dateien an vierundzwanzig kleinen und mittelgroßen deutschen Repositorien ab [Mül15, S. 32]. Dabei wurden über 150.000 PDF-Dateien ausgewertet. Festgestellt wurde, dass nur knapp acht Prozent der Dokumente als PDF/A Versionen vorlagen, stattdessen waren über fünfzig Prozent der PDF in den Versionen 1.4 und 1.6 archiviert (ebd., S. 59). Bei der Massen-Validierung der Dateien wurde festgestellt, dass über 8.500 der PDF/A-Dateien nicht valide waren und damit ungeeignet für die LZA (ebd., S. 63).

Vierkant und Kindling stellten in ihrem Artikel den 2014 Census der Open Access Repositorien vor. Dort wurden auf Basis einer "qualitativen Inhaltsanalyse der Webseiten der Repositorien" [VK14, S. 52], sowie der Abfrage von Metadaten und einer Umfrage, Daten von 173 Open Access Repositorien in Deutschland, Österreich und der Schweiz erhoben (ebd., S. 54). Die Kriterien für die Inhaltsanalyse wurden für den Census 2012 erstellt und für 2014 ergänzt. Darunter Daten zur Software, DINI-Zertifikat und Art des Trägers. Unter den zehn Kernaussagen für Deutschland sind u.a., dass Träger meist Universitäten sind, überwiegend die Softwareplattform OPUS eingesetzt wird, ein Drittel DINI zertifiziert ist und ebenfalls ein Drittel die Berliner Erklärung unterzeichnet hat (ebd., S. 63f).

²⁹<https://public.ccsds.org/pubs/650x0m2.pdf>

³⁰http://www.dlorg.eu/uploads/Booklets/booklet21x21_nutshell_web.pdf

Eine Übersicht der Auswertungen des "2014 Census on Open Access Repositories in Germany, Austria and Switzerland" ist Online verfügbar³¹ und enthält weitere vergleichende Arbeiten zu Open Access Repositorien von Kindling und Vierkant.

Bader analysierte die Daten des 2014 Census (siehe oben) für sechzehn Schweizer und fünf österreichische Repositorien. Als Kriterien wurden unter anderem Metadatenqualität, Software, Langzeitarchivierung und Lizenzen genutzt [Bad16, S. 113]. Die Autorin kam zu dem Ergebnis, dass die zu Deutschland vergleichsweise niedrige Anzahl an Repositorien in den Alpenländern eine höhere Anzahl an Datensätzen vorhält, Träger oft Universitäten und Forschungseinrichtungen sind, keins der untersuchten Repositorien zertifiziert ist, die Metadatenqualität eher gering ist und vorwiegend die Softwareplattform ePrints benutzt wird (ebd., S. 130f.).

In ihrer Masterarbeit untersuchte Bader die Situation von Open Access Repositorien in Skandinavien. Grundlage war eine quantitative Analyse von 116 Repositorien anhand von Daten aus OpenDOAR und eine inhaltliche Analyse der Webseiten der Repositorien [Bad17, S. 19]. Zur Ergänzung erfolgten Experteninterviews. Der Kriterienkatalog der quantitativen Untersuchung orientiert sich an [VK14] und dem "Open Access Repository Ranking 2015". Automatisch abgefragt wurden Daten zu Aspekten wie Repositorientyp, Art des Trägers und verwendete Software. Die Analyse der OpenDOAR Daten ergab z.B. dass die OA- Repositorien meist institutionelle Repositorien sind (ebd., S. 45) und ein Großteil von Universitäten getragen wird. Die Verteilung der Repositoriensoftware ist sehr heterogen, zeigt aber eine Tendenz zu DSpace (ebd., S. 46). Die Analyse der Webseiten war Grundlage für sog. "Backgroundchecks" (ebd., S. 7f.) zur Darstellung der Open Access Situation in den jeweiligen Ländern. Die Experteninterviews konnten Gemeinsamkeiten und Unterschiede für die skandinavischen Länder feststellen, darunter allen gemeinsam die Existenz einer nationalen Open Access Strategie (ebd., S. 71), und nationalen Open Access Metadatenportalen (ebd., S. 72). Unterschiedlich sind z.B. die präferierten Open Access Wege (grün oder gold) (ebd., S. 72).

Vergleich von Forschungsdatenrepositorien ohne einen Kriterienkatalog

Vlaeminck und Wagner untersuchten 46 europäische Forschungsdaten- und Datenservicezentren im Bereich der Sozialwissenschaften hinsichtlich ihrer Dienstleistungen zum FDM. Die Arbeit wurde in deutsch [VW13] und englisch publiziert [VW14]. Dabei wurde zu Beginn auf den Webseiten der Zentren recherchiert (in der engl. Version auch "Desktop-Research" genannt), gefolgt von einem Online-Fragebogen [VW13, S. 6]. Die Recherche auf den Webseiten der FDZ nach den FDM- Dienstleistungen war nicht ausreichend um die Forschungsfragen zu beantworten (ebd., S. 7). Der daraufhin erarbeitete Fragebogen wurde von 22 der 46 datenhaltenden Institutionen beantwortet (ebd., S. 7). Davon akzeptieren über 70% externe Forschungsdaten (ebd., S. 8) und nutzen das disziplinspezifische Metadatenschema DDI (ebd., S. 11). Etwas über die Hälfte der FDZ haben persistente Identifikatoren implementiert und 65% unterstützen ihre Nutzerx bei der Erstellung von Metadaten (ebd., S. 11f.). Der Einsatz von Maßnahmen zur Sicherung der Langzeitverfügbarkeit von FD wird von 80% der FDZ angegeben (ebd., S. 13). Insgesamt werden Forschungsdatenzentren als adäquate Partner für die Ablage von FD identifiziert.

Blumesberger (2015) präsentiert in ihrem Artikel die Zwischenergebnisse des im Rahmen des "e-Infrastructures Austria" Projektes erstellten Vergleiches des Zustandes von Metadaten an

³¹<https://oanetzwerk.wordpress.com/2015/09/07/forschungsdaten-und-auswertungen-zum-2014%2Dcensus-on-open-access-repositories-in-germany-austria-and-switzerland/>

FDR in österreichischen Institutionen. Als Methode zur Datenerhebung wurde von zuständigen Mitarbeiterx an den ausgewählten Repositorien ein Fragenbogen ausgefüllt, dies wurden gleichzeitig telefonisch von den Projektmitgliedern betreut [Blu15, S. 519]. Der Fragebogen umfasste 21 Fragen zu Metadatenschemata, Qualitätsmanagement, Barrierefreiheit, usw. Als vorläufiges Ergebnis wurde festgestellt, dass die praktische Umsetzung für alle betrachteten Themen noch nicht ausgereift ist. Dies betrifft Gebiete wie technischen Kenntnisse, Bedeutung von Metadaten für Forscherx und Nachnutzerx (ebd., S. 519) aber besonders die Barrierefreiheit (ebd., S. 521). Es wird vorgeschlagen Metadaten inhaltlich und formal niedrigschwellig zu gestalten und selbst zum Abbau von Barrieren zu nutzen (ebd., S. 523). Die vollständigen Ergebnisse der Umfrage wurden 2016 als Deliverable des Projektes verfügbar gemacht [BPZ16] und 2017 als Artikel publiziert [BZ17]. Beide haben jedoch weniger vergleichenden Charakter als der erste Artikel.

Johnston et al. (2017) verglichen die FDM-Dienstleistungen von sechs FDR. Ziel des übergeordneten Projektes "Data Curation Network" war eine Bündelung personeller Ressourcen für die Kuratierung aller anfallenden Daten in den FDR. Dafür verfasste je ein Mitglied des Projektteams eine strukturierte Zusammenfassung der angebotenen Dienstleistungen für das instituts-eigene FDR [Joh+17, S. 2 f.]. Untersucht wurden vier Themenbereiche: FD-Dienstleistungen, eingesetzte Technologien, Policies und personelle Ausstattung (ebd., S. 3). Die Autorx kamen zu dem Schluß, dass die Serviceziele kompatibel sind, diverse Workflows grundlegend gleich verlaufen (darunter DOI-Vergabe, nutzerseitiger Datenupload), die FDR-Policies einander ähneln, sowie dass alle personell unterbesetzt sind (ebd., S. 22).

Vergleich von Forschungsdatenrepositorien anhand eines Kriterienkatalogs

In der Masterarbeit von Simukovic (2012) wurden Enhanced Publications, also Artikel die mit FD angereichert werden, in je fünf FDR aus fünf Wissenschaftsdisziplinen untersucht. Ziel war die Identifikation von Gemeinsamkeiten und Unterschieden im Umgang mit Enhanced Publications [Sim12, S. 6]. Das Vorgehen wurde in drei Teile geteilt, von denen hier vor allem Teil 2 relevant ist: Hier wurden FDR anhand eines Kriterienkataloges (basierend auf Fachliteratur) hinsichtlich Zugänglichkeit, persistenter Identifizierung, Dokumentation der FD usw. untersucht (ebd., S. 5). Der Katalog wurde beschränkt auf Kriterien, die die Verbindung von FD und Publikationen betreffen. Die Auswahl der Repositorien geschah anhand von vier Referenzlisten (darunter openDOAR und der re3data Vorgänger Databib). Auswahlkriterien waren Angehörigkeit zu den gewählten Fachdisziplinen und Erwähnung in mindestens zwei der Listen. Die besten Bedingungen für die Umsetzung von Enhanced Publications wurden in der Archäologie vorgefunden (ebd., S. 58). Als Ideallösung wurde eine Verlinkung von der Publikation zum Datensatz in einem Repository identifiziert (ebd., S. 4).

Koudela, Köhler und Müller-Pfefferkorn erstellten ihren Projektbericht zu "RADIESCHEN"³² (2013) auf Basis eines Methodenmixes. Darunter Interviews, Analyse von Metadaten aus re3data, Recherche in Veröffentlichungen und auf Webseiten von FDR und Auswertung des Nutzerfeedback aus Workshops [KKM13, S. 3]. Das Ziel war es, die wichtigsten Aspekte der Technik und ihrer Einsatzmöglichkeiten im wissenschaftlichen Alltag zu charakterisieren. In der Analyse der 116 FDR die Anfang 2013 in re3data verzeichnet waren, wurden technische Aspekte abgefragt. Darunter die Repositorien Software, welches System für PID benutzt wurde und die

³²<http://www.forschungsdaten.org/index.php/Radieschen>

Art des Inhalts (ebd., S. 22). Ergebnis der Analyse war, dass die meisten FDR selbstentwickelte Repositoriensoftware einsetzten und die Eigenentwicklungen die größte Variation an verschiedenen Arten von FD aufwiesen (ebd., S. 22). Aus den Workshops wurde das Fazit gezogen, dass eine systematische Erfassung der technischen Lösungen bestehender FDR als Guideline aber nicht als Vorgabe gewünscht wurde. Die Autorx kommen zu dem Schluss, dass viele Arbeitsschritte des FDM grundlegend disziplinunabhängig sind: Disziplinübergreifende Werkzeuge für Metadatenerzeugung, Datentransfer, (Langzeit-)Archivierung, webbasierten Zugriff etc. sind möglich. Als eher heterogen und schwer zu verallgemeinern werden die Gebiete Datenerfassung, Qualitätskontrolle und disziplinspezifische Metadaten (ebd., S. 25) eingeschätzt. Hou, Thompson und Palmer untersuchten 38 öffentlich zugängliche FDR aus den Klima- und Atmosphärenwissenschaften anhand eines Kataloges mit 55 Kriterien. Der Kriterienkatalog wurde aus Fachliteratur zusammengestellt. Methodisch wurde eine Inhaltsanalyse der Webseiten und dort gesetzten Verlinkungen durchgeführt [HTP14, S. 2]. Die Zusammenstellung der zu untersuchenden Repositorien geschah anhand von thematischen Schlagworten in den Repositorienverzeichnissen DataBib und EarthCube. Ziel war die Erstellung von Profilen aller Repositorien zu deren Angeboten und Dienstleistungen um für die Zielgruppen (Forscherx, Bibliothekarx usw.) eine Vergleichbarkeit herzustellen. Die Analyse der Webseiten fand vor allem Aussagen zur Finanzierung und Organisationsform, daneben oft auch zu Datenformaten und Nutzungsbedingungen. Einige der Webseiten gaben Details zum Metadatenschema, Datensicherheit, usw. Selten wurden z.B. Informationen zu Zertifizierung und Langzeitarchivierung angegeben (ebd., S. 3). Ergebnisse der Analyse waren u.a., dass über die Hälfte der FDR eigene Metadaten verwenden, nur wenige eine Richtlinie zur Bestandserhaltung nennen und die meisten bestimmte Datenformate für den Ingest bevorzugen (ebd., S. 3). Die Autorx kamen zu dem Schluß, dass Zertifizierungen dabei helfen können das Vertrauen der Nutzerx zu gewinnen und die konsistente Anwendung von gemeinsamen Standards und Richtlinien nötig ist (ebd., S. 4).

Gómez, Méndez und Hernández-Pérez publizierten 2016 die Ergebnisse ihrer Untersuchung von FDR in den Sozial- und Geisteswissenschaften. Ziel war es die disziplinären FDR und deren Arten von FD zu identifizieren, sowie die meistgenutzten Metadaten schemata herauszuarbeiten. Im ersten Schritte wurden 397 thematisch passende Repositorien aus re3data ausgewählt, deren Metadaten per API abgefragt und die Ergebnisse quantitativ analysiert. Das Ergebnis war eine Verteilung der behandelten Themen der FDR in re3data. Bsp.: 155 FDR gaben an, FD aus den Sozialwissenschaften zu enthalten [GMH16, S. 549]. Als Metadatenstandard werden Dublin Core und DDI am häufigsten genannt, insgesamt haben aber nur 25% der FDR ein Schema hinterlegt (ebd., S. 551). In einem zweiten Schritt wurden die, auf den Webseiten angegebenen, Metadaten schemata von sechs der FDR qualitativ untersucht, je drei Geistes- und drei Sozialwissenschaftliche. Die sozialwissenschaftlichen FDR zeigten einen Trend zum DDI Schema. Für die geisteswissenschaftlichen FDR wurden multiple Schemata festgestellt, darunter DC (ebd., S. 551). Abschließend wird re3data als "registry par excellence" (ebd., S. 552) für das Auffinden von FDR bezeichnet, dessen Metadaten schema und Dienstleistung dennoch weiterentwickelt werden muss.

Newell (2017) untersuchte in ihrer Masterarbeit das FDM an 20 Hochschulen; die Auswahl erfolgt dabei an den Top 5 im Shanghai-Ranking aus Australien, Deutschland, UK und den USA. Dreizehn Kriterien wurden aus der Fachliteratur erarbeitet [New17, S. 10]. Nur das technische Kriterium "ForschungsdatenRepository" ist in diesem Kontext relevant, denn dort wurden Dateiformate, Lizenzierungsmodelle, Identifikatoren usw. untersucht. Die Untersuchung ergab, dass die FDR kaum länderspezifische Merkmale aufweisen (ebd., S. 66). Unterschiede wurden in der Verteilung von eigens konzipierten FDR und Repositorien die FD zusätzlich akzeptieren

deutlich: Australien weist überwiegend genuine FDR auf. In den UK finden sich ausschließlich Repositorien die Publikationen und FD gemeinsam aufnehmen. Diverse Aspekte der FDR werden mehrheitlich als länderunabhängig bewertet: Es werden archivfreundliche Formate empfohlen, die Archivierung ist meist kostenfrei, Forscherx können FD und MD selbst hochladen, den Zugriff über Creative Commons Lizenzen regeln und die FD über DOI persistent verlinken. Als weitere Gemeinsamkeit wird die Empfehlung von re3data zur Auswahl eines passenden FDR auf den Webseiten der Institutionen genannt (ebd., S. 66ff.).

Franzke untersuchte 2017 sieben nationale und internationale FDR in den digitalen Geisteswissenschaften. Als Auswahlkriterien werden "hohe Rezeption" und eine "explizite Ausrichtung des Repositoriums auf Geisteswissenschaften und DH" [Fra17, S. 18] genannt. Als Basis der Bewertung dienten die Webseiten der Repositorien und deren Einträge in re3data. Untersucht wurden sieben aus der Fachliteratur extrahierte Kriterien, z.B. erlangte Zertifizierungen und die Verwendung von persistenten Identifikatoren (ebd., S. 24). Die Autorin identifizierte drei vertrauenswürdige Repositorien für die DH: das Hamburger Zentrum für Sprachkorpora, das niederländische EASY-DANS und das GAMS der Universität Graz (ebd., S. 20). Empfohlen wird den FDR u.a. die Benutzung von Standards, transparente und zielgruppengerechte Informationen auf den Webseiten und Schulungen zum FDM (ebd., S. 21).

Kindling et al. veröffentlichten 2017 eine statistische Analyse der 1381 re3data Einträge, die bis inklusive 2015 erstellt wurden. Ziel war es den State of the Art von FDR für den Zeitraum zu erfassen und einen Überblick über die erfassten Angaben (Typ des FDR, Inhalte, Träger usw.) zu gewinnen [Kin+17, S. 3]. Das Metadatenschema von re3data sieht bis zu 42 verschiedene Angaben für jeden Eintrag vor, darunter allgemeine Informationen, Details zum Träger, rechtliche und technische Aspekte (ebd., S. 2). Die Daten wurden per SQL Query heruntergeladen und per Python Script normalisiert um die statistische Auswertung in SPSS zu ermöglichen. Das Resultat waren heterogene aber disziplinabhängige Ergebnisse (ebd., S. 15). Generell wurde festgestellt, dass in den FDR zu wenige der relevanten Standards (Persistente Identifier, APIs) genutzt werden. Es wird empfohlen vor allem auf Datenebene weiter zu forschen, spezifisch genannte werden Dateiformate als Ansatzpunkt (ebd., S. 16). Abschließend werden neun Empfehlungen für Repository Manager ausgesprochen, darunter die Prüfung des eigenen re3data- Eintrages, die Verwendung von APIs, PID und Standardlizenzen und das Ablegen einer Zertifizierung (ebd., S. 17).

Einige Besonderheiten von Videos als Forschungsdaten

Die Behandlung von Forschungsdaten in einem Repository ist nicht zwingend abhängig von deren Art. Beispielsweise ist für ein Embargo irrelevant ob ein Video oder eine Tabelle gesperrt wird. Für andere Punkte hingegen ist eine besondere Behandlung von Videos möglich und wird in dieser Arbeit geprüft. Hier eine beispielhafte Übersicht von Besonderheiten im Umgang mit Video-Forschungsdaten:

K1 Veröffentlichung und Nachnutzung

Datenabgabe: Als archivfreundliche Formate für Video werden MPEG-4, OGG video und motion JPEG 2000 empfohlen [UK oJ], bzw. Matroska (.mkv) [nes16, S. 58]. Eine Begrenzung der akzeptierten Formate wird empfohlen (ebd.). Mit Stand 2012 wurde das Containerformat MXF empfohlen, in Benutzung waren außerdem MOV, AVI und JPEG 2000 [Wri12, S. 17].

Datenintegrität: Empfohlen wird eine Eingangskontrolle, die mindestens Lesbarkeit des Videos, Vollständigkeit, technische Fehlerfreiheit u.ä. prüft [nes16, S. 51].

Metadaten: FDR ermöglichen Angaben für Rechte aus verwendeter Musik in den Metadaten

[Ing17]. Für die Beschreibung von Videos werden spezielle Metadaten-Schemata wie PBCore³³ oder videoMD³⁴ angeboten. Die Metadaten-Schemata DC und METS sind laut [RB10] geeignet für die Beschreibung von Bewegtbildern. Eine große Übersicht von Metadaten-Schemata, auch für "Moving Images", ist Online verfügbar³⁵. Die nestor-AG Media empfiehlt außerdem PREMIS [nes16, S. 52].

Zitierfähigkeit: Es besteht die Möglichkeit einen bestimmten Zeitpunkt in einem Video zu zitieren ([Ing17], [Wri12, S. 19]), analog der Angabe von Seiten in textuellen Quellen. Beispielsweise durch Nutzung der Media Fragments URI des W3C³⁶. Der British Universities Film & Video Council veröffentlichte bereits 2013 die "Audiovisual citation guidelines". Im Dezember 2017 erschien die zweite Auflage³⁷.

Nutzungsstatistik: Es gibt z.B. eine eigene Nutzungsstatistik für Anzahl der Aufrufe ("Views"), um diese von den Aufrufen der Landingpages zu differenzieren, die dazu nötigen Hintergründe liefert z.B. der COUNTER Code of Practice³⁸.

K2 Zugriff und Auffindbarkeit

Recherchemöglichkeiten: Video Retrieval bietet eine große Anzahl von Ansätzen um eine Fülle an Daten und Informationen aus Videos zu extrahieren, vgl. dazu Übersichtsarbeiten wie [Rai16], [Eid12], [Wri12, S. 19]. Beispiele für spezielle Suchmöglichkeiten in Videos liefert z.B. das TIB AV-Portal³⁹.

K3 Transparenz und Qualitätssicherung

Sicherstellung langfristiger Zugänglichkeit: Der schnelle Wandel an Hard- und Softwareformaten für alle multimedia Formate ist eine besondere Herausforderung, genauso wie verlustreiche und den Bitstream beschädigende Algorithmen zur Reduzierung großer Datenmengen [Deg14, S. 48]. Migration in andere Formate ist für Videos besonders kostenintensiv [Ing17], wird aber dennoch empfohlen [nes16, S. 54]. Der Codec muss für jedes Video bekannt sein (ebd., S. 54), empfohlen wird FFV1 (ebd., S. 58). Kapitel 5 aus Wright [Wri12] stellt diverse weitere Details zur LZA von Videos bereit.

K4 Nutzererlebnis und Support

Usability: Nachnutzende Personen profitieren z.B. von Annotationsmöglichkeiten für Videos [Ing17], Jisc bietet das Tool "Clipper"⁴⁰. Barrierefreiheit für Videos wird z.B. durch Transkripte, Untertitel und inhaltliche Beschreibungen unterstützt [Cal+09].

Darstellung der Datensätze: Streaming ist die derzeit Online genutzte Vorschaufunktion für Videos.

K5 Organisation und Infrastruktur

Personalausstattung und -qualifizierung: Enthält Angaben zu Mitarbeiter:innen, die sich auf die Behandlung von Videos spezialisieren oder deren Aufgabe jegliche Behandlung von Videos im FDR beinhaltet.

Forschungs- und Entwicklungstätigkeiten: Teilnahme an Projekten in denen Videos zentral oder wenigstens relevant sind.

Technische Infrastruktur: Erwähnt besondere Maßnahmen die für Videos getroffen werden. Problematik von Backups für große Datenmengen wird thematisiert [Ing17].

³³<http://pbcore.org/>

³⁴<https://www.loc.gov/standards/amdvmd/>

³⁵<http://jennriley.com/metadatamap/seeingstandards.pdf>

³⁶<https://www.w3.org/TR/media-frags/>

³⁷<http://bufvc.ac.uk/projects-research/avcitation>

³⁸https://www.projectcounter.org/wp-content/uploads/2017/07/Friendly_Guide_Providers_20170706-1.pdf

³⁹<https://www.tib.eu/de/recherchieren-entdecken/audiovisuelle-medien/>

⁴⁰<http://blog.clippertube.com/>

K6 Gebühren und Finanzierung

Zu Gebühren für Datenabgabe und -nutzung und zur Finanzierung des Betriebs sind keine Besonderheiten von Videodaten bekannt.

Arbeiten die sich mit dem Vergleich von Forschungsdatenrepositorien beschäftigen und einen Schwerpunkt auf audiovisuelle Forschungsdaten (Ton, Video oder sonstige bewegte Bilder) legen, wurden im Rahmen der Recherche nicht gefunden.

4 Vorstellung der Forschungsdatenrepositorien

Stanford Digital Repository

Das Stanford Digital Repository (kurz: SDR) ist das institutionelle Repository der Stanford University in Kalifornien, USA. Mit Stand 22.12.17 werden 258.883 Datensätze angezeigt, davon 4.536 bzw. 2% Video. Diverse Typen von Daten werden vorgehalten, die häufigst verwendeten sind "Image" und "Map" (siehe Screenshot SDR004_Typen von Daten).

Das Repository wurde 2006 gegründet [Cra11, Folie 4] und akzeptiert seit 2012 auch Forschungsdaten [Hod14]. Betreut werden laut re3data die Disziplinen Geistes-, Sozial-, Natur- und Ingenieurwissenschaften [re317e].

Im Bibliotheksblog wird unter dem Topic "Stanford Digital Repository"⁴¹ über das "SDR deposit of the week" und diverse andere Themen berichtet.

Phaidra

Phaidra ist ein Bestandteil des institutionellen Repositoriums der Universität Wien. Aufgelöst steht die Bezeichnung für "Permanent Hosting, Archiving and Indexing of Digital Resources and Assets" [Uni Jb]. In dieser Arbeit wird ausschließlich Phaidra betrachtet. Nicht behandelt werden die anderen Phaidra-basierten Repositorien "Phaidra Temp"⁴² und "Phaidra Sandbox", die zusammen mit Unidam⁴³ und Github das verteilte Archivierungssystem der Universität Wien bilden [BG17, S. 22].

Die zugrundeliegende Software wurde 2008 für die Universität entwickelt und wird inzwischen als "Phaidra Netzwerk" an über einem Dutzend Hochschulen in fünf Ländern eingesetzt [Uni Jc].

Alle großen wissenschaftlichen Disziplinen werden betreut [re317d] und die Typen von Daten zeigen, dass sowohl Publikationen als auch Forschungsdaten im Repository liegen. Es werden vor allem Bilder und digitalisierte Bücher eingespeist. Mit Stand 22.12.17 wurden 71,568 Datensätze verzeichnet, davon 1,148 (ca. 2%) als Video markiert (siehe Screenshot PHA001_Typen von Daten).

Zenodo

Zenodo wurde 2013 zuerst unter dem Namen "OpenAIRE Orphan Record Repository" Online gestellt und wird von CERN gehostet [Zen Jc].

Es bezeichnet sich als "catch-all" Repository, dass alle Arten von Forschungsdaten ohne Einschränkung akzeptiert [Zen Ja]. Re3data führt Zenodo in der Kategorie repositoryType als "other" [re317f].

Mit Stand 22.12.17 gibt Zenodo eine Anzahl von 312.957 Suchergebnissen aus, die Datentypen Image und Publication machen davon den größten Anteil aus. Etwa 0,2% der Ergebnisse beinhalten Videos (siehe Screenshot Zen007_AnzahlVideos).

Brown Digital Repository

Das Brown Digital Repository ist das institutionelle Repository der Brown University in Rhode Island, USA. Auch hier werden Daten aus allen großen Wissenschaftsdisziplinen akzeptiert,

⁴¹<https://library.stanford.edu/blogs/topic/stanford-digital-repository>

⁴²<https://phaidra-temp.univie.ac.at/>

⁴³<https://unidam.univie.ac.at/>

sowie alle Dateitypen [Bro Je]. Mit Stand 22.12.17 beinhaltet dies 181.254 Suchergebnisse, überwiegend bestehend aus den Datentypen text und still image. Videos machen einen Anteil von 0,1% aus (siehe Screenshot BDR009_Typen von Daten).

LSE Research Online

Der volle Name des institutionellen Repositoriums lautet "London School of Economics and Political Science Research Online" und wird mit LSERO abgekürzt.

Entsprechend der Fachrichtung des Trägers werden hier laut re3data nur Daten aus den Sozial-, Wirtschafts-, Verhaltens- und Geisteswissenschaften [re317b] vorgehalten. Mit Stand 22.12.17 werden 68.385 Suchergebnisse angezeigt (siehe Screenshot LSE004_AnzahlErgebnisse), davon 40 als Video markiert (siehe Screenshot LSE005_AnzahlVideo). Auch hier werden Publikationen gemeinsam mit Forschungsdaten abgelegt.

OceanRep GEOMAR Repository

Das institutionelle Repository des Geomar Helmholtz-Zentrum für Ozeanforschung Kiel wird als OceanRep GEOMAR Repository oder nur OceanRep bezeichnet. Das GEOMAR Institut forscht auf dem Gebiet der Meeresforschung [GEO17] bzw. marinen Naturwissenschaft [GEOoJ].

Von den aktuell 21.520 Einträgen in OceanRep sind die meisten der Kategorie Artikel zugeordnet, gefolgt von Hochschulschriften. Die 28 verzeichneten Videos machen nur 0,1% des Bestandes aus (siehe Screenshot ORE006_Typen von Daten mit Stand 22.12.17). Eigentlich handelt es sich bei OceanRep um eine Publikationsdatenbank, doch durch die Listung in re3data und im Sinne der festgelegten Definition von Forschungsdaten fällt es mit den enthaltenen Videos dennoch in die Zielgruppe und wird im Folgenden analysiert.

OceanRep wird gemeinsam von der GEOMAR-Bibliothek und vom Datenmanagementteam des Instituts betreut [GEOoJ].

Marine-Geo Digital Library

Die Marine-Geo Digital Library (kurz: MGDL) des Marine Geoscience Data System (kurz: MGDS) wird an der Columbia University gehostet. Als disziplinäres und institutionelles FDR bietet es Zugang zu "[...] marine geoscience research data acquired throughout the global oceans and adjoining continental margins." [Mar Jg]. Grundlage der MGDL waren Projekte der US-Amerikanischen National Science Foundation im Zeitraum 2003-2004 [Lam J]. Das FDR enthält mehrere sogenannte Datenportale (darunter GeoPRISMS und das Antarctic & Southern Ocean Data Portal), die jedoch eher als Sammlungen mit Communityspezifischen Sucheinstiegen zu verstehen sind. Die Portale sind Sucheinstiege in ein gemeinsames System auf einer integrierten Infrastruktur [Uni Ja]. Daher fällt die MGDL weiterhin unter das Kriterium Pull.

Mit Stand 04.01.18 werden 8642 Datensätze vorgehalten (siehe Screenshot MGD001_ErgebnisseGesamt), darunter drei Videos (siehe Screenshot MGD002_ErgebnisseVideo). Den Hauptbestandteil machen Sensor- und Messdaten aus, sowie die daraus entstandenen Datenprodukte [Mar Jd]. Das MGDS ist Teil der Interdisciplinary Earth Data Alliance (IEDA)⁴⁴.

DR-NTU Data Repository

DR-NTU (Data) ist das institutionelle FDR der Nanyang Technological University in Singapur.

⁴⁴Verzeichnet als "Marine Geology & Geophysics": <https://www.iedadata.org/data-repositories-syntheses/>

Es wurde im September 2017 in Betrieb genommen [Wen17] und im November eingeweiht [Aso17].

Mit Stand 23.12.17 enthält es 180 Datensätze, davon ein Video (siehe Screenshot NTU005_Typen von Daten). Vorrangig wird die Kategorie "unknown" angegeben, die eine große Anzahl verschiedener Dateitypen vorweist.

Die Universität beschreibt das übergeordnete Repositoryum DR-NTU als "one-stop portal" [Nan15], es ist unterteilt in DR-NTU Open Access⁴⁵, DR-NTU Restricted Access⁴⁶ und das hier betrachtete DR-NTU (Data)⁴⁷.

⁴⁵<https://dr.ntu.edu.sg/>

⁴⁶<https://repository.ntu.edu.sg/>

⁴⁷<https://researchdata.ntu.edu.sg/>

5 Analyse der Forschungsdatenrepositorien

Die Grundlage der Analyse der FDR anhand des Kriterienkataloges befindet sich in Anhang A.2, im folgenden werden ausgewählte Ergebnisse der Analyse für jedes FDR vorgestellt. Die Reihenfolge orientiert sich dabei an Steinhof (2017).

5.1. Stanford Digital Repository

K1 Veröffentlichung und Nachnutzung

Das SDR akzeptiert alle Datenformate [Sta Jh], hat aber "Best Practice File Formats"⁴⁸ zusammengestellt. Der Upload erfolgt über das "SDR Online Deposit" Formular im Webbrowser. Für Datenmengen über 10GB sollen Nutzerx Kontakt mit den Mitarbeiterx aufnehmen [Sta Je]. Pflichtfelder werden gekennzeichnet, die Vergabe von Schlagworten erfolgt frei [Fro13].

Als Metadatenstandards werden MODS und Dublin Core [Sta Jk] verwendet, außerdem wird auf FGDC, EAD und DDI [Sta Ja] verwiesen. Die Nutzerx werden bei der Erfassung und Eingabe von Metadaten durch drei Webseiten unterstützt: "Metadata Basics"⁴⁹, "Advanced Metadata"⁵⁰ und eine Liste von Metadaten-Werkzeugen⁵¹. Aktuell werden CEDAR Templates⁵² getestet [Egy+17, S. 9] getestet.

Die Dokumentation der FD kann als "related items" oder gemeinsam mit Forschungsdaten abgelegt werden [Fro13]. Die Verlinkung auf eine Datenpublikation ist über "related items" möglich (siehe Screenshot SDR002_RelatedItems).

Auf Datensatzebene wird ein Zitationsvorschlag (siehe Screenshot SDR001_PreferredCitation) aus Metdaten generiert oder vom Forscherx selbst formatiert [Fro13]. Als PID wird PURL mit DRUID genutzt [Sta Jk].

Das SDR bietet den Forscherx zur Lizenzierung alle CC- und zusätzlich die ODC- Lizenzen an [Sta Jb].

Eine direkte Verlinkung von FD und Publikationen ist vorgesehen [Sta Jf], ein Link zu Publikationen die FD nachnutzen ist nicht explizit erwähnt, aber möglich über die Related Content Funktion. Selbige ermöglicht auch die Verknüpfung zugehöriger Projekte [Fro13, Minute 9:19]. Die Erhebung von Metriken, Zitationsraten oder Altmetrics auf Datensatzebene ist bisher nicht implementiert [Sta Je].

Als Anreize für Forscherx wird die Sichtbarkeit der FD erhöht, indem sie gemeinsam mit allen anderen Bibliotheksressourcen angezeigt [Sta Jd] werden. Zusätzlich gibt es "Featured Collections" [Sta Jf] und das "SDR Deposit of the Week"⁵³ auf dem Blog der Universitätsbibliothek. ORCID's können bisher nicht eingegeben werden, aber Interesse auf Seiten des SDR besteht

⁴⁸<https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-formats>

⁴⁹<https://library.stanford.edu/research/data-management-services/data-best-practices/creating-metadata/basic-approach-metadata>

⁵⁰<https://library.stanford.edu/research/data-management-services/data-best-practices/creating-metadata/advanced-metadata>

⁵¹<https://library.stanford.edu/research/data-management-services/data-best-practices/creating-metadata/metadata-tools>

⁵²<https://metadatacenter.org/tools-training/cedar-template-tools>

⁵³Beispiel: <https://library.stanford.edu/blogs/digital-library-blog/2017/05/sdr-deposit-week-video-tutorials-4d-visualization>

[Sta Jk].

K2 Zugriff und Auffindbarkeit

Zum Upload von FD sind Universitätsangehörige berechtigt [Sta Je]. Der Zugriff kann öffentlich oder "Stanford-only" eingeschränkt werden [Sta Jb], eine Einschränkung auf Benutzergruppen ist geplant [Sta JI]. Die technische Zugriffsteuerung wird beschrieben als "[...] controlled using strict authentication policies and enterprise-level security mechanisms." [Sta Ji].

Das Auferlegen eines Embargos ist auf Datensatz- [Sta Ji] und Sammlungsebene [Fro13] für bis zu 3 Jahre möglich [Sta Jd].

Eine Suche im SDR ist über die Einschränkung auf die entsprechende Kollektion im Bibliothekskatalog "SearchWorks" möglich, dort ist das SDR seit 2013 eingebunden [Sta I3].

Search Works bietet eine freie Textsuche mit Booleschen Operatoren, Trunkierung, Wildcards usw. (siehe Screenshot SDR005_Suche). Eine Suche über Metadaten ist möglich (siehe Screenshot SDR006_SucheMD). Die Facettierung bietet diverse Kategorien (siehe Screenshot SDR007_Facettierung).

Da das SDR im suchmaschinenfreundlichen OPAC verzeichnet ist, können die FD auch extern gefunden werden, explizit genannt wird Google [Hod17].

Die Abfrage von FD ist über eine RESTful API möglich, der Upload über Schnittstelle aber noch nicht [Sta Jk]. Seit 2016 bietet das SDR eine IIIF Schnittstelle für Bilder [Sny16], an der es auch Mitentwickler ist⁵⁴.

K3 Transparenz und Qualitätssicherung

Explizit als Policy werden nur die "SDR Terms of Deposit"⁵⁵ bezeichnet. Die sonst üblichen Inhalte von Richtlinien sind verteilt auf den Hilfeseiten für Forscherx und Nutzerx, z.B. SDR Overview [Sta Ji], FAQ [Sta Je] usw.

Für die Versionierung von Datensätzen wird die Eigenentwicklung Moab verwendet [Sta Jj]. Moab unterstützt Checksums [And13, S. 29] und protokolliert Änderungen (ebd., S. 15). Die aktuellste Version eines Datensätzen wird auf der Webseite angezeigt, ältere Versionen sind auf Anfrage zugänglich [Sta Jd]. Das Löschen einzelner Dateien ist über Versionierung möglich [And13, S. 25].

Das Ziel des SDR wird beschrieben als: "[...] preserve, manage, and share research data in a secure environment for long-term citation, access, and reuse." [Sta Jc] und ist der einzige Hinweis auf Übernahme der Verantwortung für die langfristige Sicherung. Die Daten werden redundant und örtlich verteilt aufbewahrt, aber es werden keine "formal services in format migration or system emulation" [Sta Jk] angeboten. Die Aufbewahrung geschieht auf Magnetband und Hard Disk Drives [Sta Ji], dabei wird eine Kopie auf Festplatte vorgehalten und die übrigen auf LTO Magnetbändern [Sta Jj]. Zusätzlich werden die Daten regelmäßig geprüft, um ungewollte Änderungen zu erfassen [Sta Jd].

Das SDR ist Mitglied des Data Preservation Network⁵⁶ [Sta Jd].

K4 Nutzererlebnis und Support

Die Weboberfläche des SDR ist nicht Barrierefrei. Die Browser-Erweiterung aXe zeigt 18 Verstöße auf Datensatzebene an (siehe Screenshot SDR008_Accessibility-Datensatz), es bestehen vor allem Probleme mit Listenelementen, fehlender Beschriftung und niedrigem Farbkontrast. Die Anzeige der Suchergebnisse ergibt 50 Verstöße (siehe SDR009_Accessibility-Suche). Die manuelle Betrachtung ergab, dass Transkripte für Videos technisch möglich sind, aber nicht konsistent vorliegen oder durchsuchbar sind (siehe Screenshot SDR010_Transkript).

⁵⁴ <https://library.stanford.edu/projects/international-image-interoperability-framework>

⁵⁵ <https://web.stanford.edu/group/sdr/SDR-Terms-of-Deposit-v13.pdf>

⁵⁶ <https://dnp.org/>

Für Forscherx und Nutzerx werden diverse Hilfeseiten geboten⁵⁷, darunter auch Guides und Videotutorials⁵⁸. Die FAQs sind dreigeteilt in "General", "Copyright, licenses and access restrictions" und "Technical topics". Die Universitätsbibliothek bietet diverse Schulungsangebote zum FDM⁵⁹. Kontaktmöglichkeiten für Supportanfragen sind im Footer verlinkt (siehe Screenshot SDR012_Supportanfrage).

Für die Beratung zum FDM gibt es unter anderem "Data management", "Digitization" und "General consultation" Services [Sta Jd]. Hilfestellungen gibt es u.a. zum Thema "Share and preserve research data"⁶⁰ oder über die "Managing Data FAQ"⁶¹. Die "Data best practices"⁶² und eine Liste disziplinspezifischer Repositorien⁶³ ergänzen das Angebot.

K5 Organisation und Infrastruktur

Das SDR ist organisatorisch eine Abteilung der Stanford University Library [Sta Ji]. Getragen wird das FDR von der Stanford University und Stanford University Library gemeinsam [Sta Ji]. Die Bibliothek ist Gemeinnützig und bezeichnet das SDR als "[...] unencumbered by commercial interests or proprietary systems [...]" [Sta Ji].

Das Team des SDR umfasst 30 Mitarbeiter, darunter "preservation experts" aber auch Softwareentwickler [Sta Ji].

Die Universität Stanford, die Digital Public Library of America und DuraSpace führen gemeinsam das Projekt "Hydra-in-a-Box" durch, mit dem Ziel ein "feature-rich, robust, flexible digital repository" [Sta17] zu entwickeln. Die Universitätsbibliothek entwickelte außerdem das Plugin "Spotlight" zur Erstellung von Webseiten für digitale Sammlungen [Sny14]. Im Bereich des FDM war die Bibliothek beteiligt am "Manifesto for Linked Libraries"⁶⁴. Diverse weitere Forschungsprojekte sind auf der Bibliothekswebseite verzeichnet⁶⁵.

Die technische Grundlage des SDR ist Hydra, welches als "complete solution composed of open source technologies" [Sta Jj] beschrieben wird. Darin enthalten sind Fedora (Repositoriensoftware), Solr (Suche), Blacklight (Facettierte Suche) und Moab (Versionierung). Für weitere technische Details wird der Kontakt zu SDR Mitarbeitern angeboten (ebd.).

K6 Gebühren und Finanzierung

Das SDR ist dabei ein Gebührenmodell zu entwickeln [Sta Jg]. Mit aktuellem Stand (Dezember 2017) werden keine Kosten angegeben.

5.2. Phaidra

K1 Veröffentlichung und Nachnutzung

Phaidra hat eine Liste mit empfohlenen Formaten⁶⁶ für den Ingest, akzeptiert aber alle Arten von Ressourcen [Abt Jg], darunter Text, Video und Audio [Abt Jf].

⁵⁷<http://library.stanford.edu/research/stanford-digital-repository>

⁵⁸<http://library.stanford.edu/research/stanford-digital-repository/depositor-services/online-deposit>

⁵⁹<https://library.stanford.edu/research/data-management-services/consulting-training%2Dand-other-services>

⁶⁰<https://library.stanford.edu/research/data-management-services/share-and-preserve%2Dresearch-data>

⁶¹<https://library.stanford.edu/research/data-management-services/faq/managing-data-faq>

⁶²<https://library.stanford.edu/research/data-management-services/data-best-practices>

⁶³<https://library.stanford.edu/research/data-management-services/share-and-preserve%2Dresearch-data/domain-specific-data-repositories>

⁶⁴<https://library.stanford.edu/projects/linked-data>

⁶⁵<https://library.stanford.edu/projects>

⁶⁶<https://backend.univie.ac.at/index.php?id=23004&L=0>

Für den Upload von FD werden drei Optionen geboten: "[...] Einzelupload, Bulkupload in Zusammenarbeit mit dem ZID oder Nutzung des PhaidralImporters." [Blu16, S.4]. Ganguly präzisiert, dass für den Upload en bloc "Docked Applications" [Gan17, S. 111] notwendig sind. Der PhaidralImporter⁶⁷ ist als Download in Phaidra verfügbar.

Wie in den erweiterten Metadaten deutlich wird (siehe z.B. Screenshot PHA005_Klassifikation), werden die Metadaten in 9 Gruppen organisiert, darunter Allgemeine und Technische Angaben, sowie Rechte und Lizenzen [Abt08, S. 24]. Das Metadatenschema basiert auf Dublin Core [Abt Jg] und einer lokal angepassten Version der LOM⁶⁸ [Abt Jf].

Bei der Eingabe der Metadaten werden Forscherx durch die erwähnten Docked Applications unterstützt, dabei handelt es sich um eine "Applikation, die sich über die Schnittstelle mit Phaidra verbindet [...]" [Gan17, S. 109]. Erfolgt die Eingabe manuell, werden Forscherx auf Pflichteingaben hingewiesen, dazu gehören Titel des Datensatzes und Lizenzangaben [Abt Jc]. Es stehen "umfassende Werkzeuge für die formale und inhaltliche Erschließung" [Abt Jf] bereit, als Beispiele werden u.a. die Österreichische Systematik der Wissenschaftszweige, der Getty Thesaurus of Geographic Names und der multilinguale Thesaurus Eurovoc genannt (ebd.). In den erweiterten Metadaten wird die genutzte Klassifikation angezeigt (siehe Screenshot PHA005_Klassifikation). Zur Arbeitserleichterung können Forscherx eine Metadatenvorlage erstellen und damit einmal eingegebene Metadaten für die nächste Eingabe wieder verwenden [Abt08, S. 25].

Blumesberger und Ganguly starteten den Aufruf ein "Netzwerk nachhaltige Softwareentwicklung" aufzubauen, dass sich mit dem Verbleib von Software aus Forschungsprojekten beschäftigt [Abt Jd]. Es gibt jedoch keine expliziten Aussagen zur Archivierung von FD-relevanter Software in Phaidra.

Für die persistente Identifikation von Webseiten (z.B. Landingpages) nutzt die Universität Wien das Handle System, sowie DOI für Forschungsdaten [Abt Jg].

Die allgemeinen Nutzungsbedingungen von FD werden in der Policy festgelegt [Abt Jf], spezifisch werden die Creative-Commons-Lizenzen empfohlen. Eine Liste der CC Lizenzen ist für Nutzerx auf der Webseite verfügbar⁶⁹.

Zur Lizenzierung der Metadaten in Phaidra gibt es keine explizite Aussage, sie werden lediglich als "öffentlich verfügbar" [Abt Jg] beschrieben.

Publikationen die auf den FD basieren können in "u:scholar", einem anderen Bestandteil des institutionellen Repositoriums, hochgeladen werden. "u:scholar" ist eine Docked Application und die dort hochgeladenen Publikationen werden in Phaidra gespeichert [Abt Je]. An anderer Stelle wird erwähnt, dass alle Arten von Datensätzen im Phaidra System miteinander verbunden sind [Abt Jg], dies schließt implizit Publikationen und Forschungsdaten mit ein.

Die Anzahl der Aufrufe der Landingpages wird auf Datensatzebene angezeigt (siehe Screenshot PHA002_AnzeigeVideo), in den Nutzungsbedingungen wird darauf hingewiesen, dass u.a. Download-Statistiken aus Log-Dateien erhoben werden [Ami J]. Außer einem Hinweis auf Anonymisierung werden keine weiteren Aussagen zur Erhebung der Nutzungsstatistik getroffen.

K2 Zugriff und Auffindbarkeit

Zum Upload berechtigt sind Universitätsangehörige, inkl. Studierenden [Abt Jg]. Datengeberx werden als aktive Verwender bezeichnet, während Nachnutzerx der Forschungsdaten als passive Verwender des FDR eingestuft werden [Abt Ji]. Die Steuerung des Zugriffs für den Upload geschieht über das u:account⁷⁰ System der Universität Wien. Die hochladende Person kann

⁶⁷https://phaidra.univie.ac.at/detail_object/o:471960

⁶⁸Weitere Informationen zu Learning Object Metadata: http://129.115.100.158/txlor/docs/IEEE_LOM_1484_12_1_v1_Final_Draft.pdf

⁶⁹<https://datamanagement.univie.ac.at/rdm/juristisches/lizenzmodelle/>

⁷⁰<http://zid.univie.ac.at/uaccount-benutzungsordnung/>

die FD öffentlich zur Verfügung stellen oder auf einen Nutzerkreis beschränken, dafür gibt es vom FDR festgelegte Nutzergruppen [Gan17, S. 109].

Phaidra bietet eine Embargofunktion, dessen Frist vom Forscherx festgelegt wird [Abt Jf]. Nach Ablauf des Embargos schaltet das System den Datensatz ohne manuellen Input frei [Gan17, S. 109].

Das FDR bietet nur eine Einschlitzsuche, welche die Metadaten durchsucht. Ganguly erwähnt außerdem eine Expertensuche [Gan17, S. 111], die aber auf der Webseite zumindest nicht für anonyme Nutzerx erreichbar ist (siehe Screenshot PHA001_Typen von Daten). Durch das Fehlen einer Suchhilfe ist der Umfang der Suchfunktionen nicht explizit dargestellt, Stichprobentests für die Suche mit Wildcards und Trunkierung (siehe Screenshot PHA010_Suchfunktionen-Trunkierung-Wildcard) und Boolesche Operatoren (siehe Screenshot PHA011_Suchfunktionen-Bool) scheinen erfolgreich. Zusätzlich sind "metadatenbasierte Suchstrategien" [Abt Jf] möglich, sowie eine facettrierte Suche nach Objekttypen, Lizenzen usw. (siehe Screenshot PHA001_Typen von Daten). Zu besserer Findbarkeit können die Metadaten durch externe Suchmaschinen indexiert werden [Ami J].

Phaidra bietet aktuell REST und OAI-PMH Schnittstellen an, eine LTI-Schnittstelle⁷¹ ist in Planung [Gan17, S. 111]. Die Dokumentation der phaidra-api für Ingest und Download von Daten ist über github erhältlich⁷², über diese interagieren auch die Applikationen (im Phaidra Kontext "Docked Applications" genannt) mit dem Datenbestand [Gan17, S. 109].

K3 Transparenz und Qualitätssicherung

Phaidra bietet den Nutzerx eine explizit benannte Policy⁷³, sowie im Footer verlinkte Nutzungsbedingungen⁷⁴. Die Universität Wien hat außerdem eine OA Policy⁷⁵.

Im Metadatenschema ist ein Feld zum Peer Review Status des Objektes vorgesehen (siehe Screenshot PHA006_PeerReview), weitere Angaben zur Datenvalidierung werden nicht gemacht.

Die Versionierung von Datensätzen nach Veröffentlichung ist möglich. Die alte Version behält dabei ihren PID und wird mit der Neuen verknüpft [Abt Jf]. Die Versionsnummer wird auf Datensatzebene angezeigt (siehe Screenshot PHA002_AnzeigeVideo).

Die Löschung von Datensätzen ist grundlegend nicht vorgesehen [Abt08, S. 19], aber möglich wenn eine Sperrung der Daten nicht ausreicht [Ami J]. Als Grund für eine Löschung werden vor allem Rechtsverletzungen angeführt (ebd.).

Die Verantwortung für die langfristige Zugänglichkeit wird in den Nutzungsbedingungen übernommen [Ami J]. Die Sicherheitsmaßnahmen orientieren sich dabei an OAI Richtlinien, diese werden aber auch "weiterentwickelt" [Abt Jf]. Spezifisch wird Datenmigration auf drei Ebenen betrieben: Bitstreampreservation für Daten und physische Medien, stetige Überführung in aktuelle Formate und die Prüfung von Semantik und Logik nach jeder Migration [BG17]. Die Sicherung wird für 5 Jahre ab Upload garantiert [Ami J]. Statt einer Nachfolgeregelung ist vorgesehen, dass die Daten sofern möglich erhalten werden und "auf Verlangen" an den hochladenden Nutzerx übergeben werden [Ami J].

Phaidra ist nicht zertifiziert.

K4 Nutzererlebnis und Support

Die Nutzerxfreundlichkeit bei der Eingabe von Metadaten und Suche nach Datensätzen wird von Ganguly eher gering bewertet [Gan17, S. 111]. Die Suchoberfläche ist in Deutsch, Englisch

⁷¹Mehr zu Learning Tools Interoperability: <http://www.imsglobal.org/activity/learning-tools-interoperability>

⁷²<https://github.com/phaidra/phaidra-api/wiki/Documentation>

⁷³<http://datamanagement.univie.ac.at/ueber-phaidra/policy-von-phaidra/>

⁷⁴https://phaidra.univie.ac.at/terms_of_use/show_terms_of_use

⁷⁵<https://openaccess.univie.ac.at/open-access/oa-policy-der-uni-wien/>

und Italienisch verfügbar, mit Stand 2012 wurde außerdem Serbisch angegeben [Abt Jf]. Phaidras Webseite ist nicht barrierefrei. Die Browsererweiterung aXe findet 11 Verstöße gegen die Barrierefreiheit nach WCAG 2.0 für die Anzeige der Suchergebnisse (siehe Screenshot PHA012_Accessibility-Suche) und 8 Verstöße auf Datensatzebene (siehe Screenshot PHA013_Accessibility-Datensatz). Die Probleme bestehen bei beiden vor allem in fehlendem Farbkontrast, aber auch fehlender Beschriftung (Buttons, Form-tags, IDs). Abgesehen davon zeigt die Analyse der Webseite, dass die Beschreibung von Videos in den Metadaten per Freitext umgesetzt werden kann (siehe Screenshot PHA004_Accessibility), dies ist aber optional und nicht normiert. Die Stichprobenprüfung für jeweils 3 Datensätze aus den Typen Video, Bild, Audio stellte in keinem Fall Textalternativen fest. Die Suche ist nicht mit Tastatur allein bedienbar. Die Schlitzsuche hat eine Eingabehilfe in Form von Suchvorschlägen (siehe Screenshot PHA014_Accessibility-Suchvorschläge).

Zur Datenvisualisierung wird angegeben, dass "mehrere Möglichkeiten zur Verfügung" stehen [Abt Jb], dies wird für passive Anwenderx nicht weiter spezifiziert.

Für audiovisuelle Forschungsdaten besteht eine Streamingoption (siehe Screenshot PHA002_AnzeigeVideo), die die Ansicht der FD auf der Webseite ermöglicht.

Eine explizite Hilfeseite oder allgemeine FAQs fehlen, stattdessen werden Informationsseiten zu Einzelthemen wie Technik⁷⁶ oder Zugang⁷⁷ angeboten. Die Webseite bietet eine Übersicht zu aktuellen Schulungen und Workshops⁷⁸, die von Mitarbeiterx der Universitätsbibliothek durchgeführt werden. Auf Anfrage werden auch Einzelschulungen angeboten [Abt Jf].

Für allgemeine und technische Anfragen wird auf e-Mail Support und den Help Desk des zentralen Informatikdienstes der Universität verwiesen [Abt Jf].

Die Beratung zum FDM wird von der Abteilung Phaidra-UNIDAM selbst durchgeführt und beinhaltet u.a. Metadateneingabe und die Erstellung von Datenmanagementplänen [Abt Jc], z.B. durch Guidelines⁷⁹. Aktive Hilfe wird z.B. bei der Antragsstellung und Durchführung von Projekten oder im Ingest angeboten [BG17, S. 23]. Zusätzlich gibt es für Universitätsangehörige eine Sandboxinstanz, also eine eigenständige Phaidra-Umgebung mit Testdaten [Abt Je].

K5 Organisation und Infrastruktur

Die IT-Abteilung der Universität betreut das FDR technisch und die Bibliothek betreut es inhaltlich. Geleitet wird das Projekt ebenfalls durch die Bibliothek [Abt Jf]. Die Abteilung Phaidra - UNIDAM⁸⁰ ist die zuständige Einheit in der Universitätsbibliothek.

Auf den Webseiten der Abteilung werden nur 3 Mitarbeiterx genannt⁸¹. Im allgemeinen Mitarbeiterverzeichnis gibt es keine weiteren Hinweise, ebensowenig wie zur Qualifizierung des Personals.

Die Phaidra Infrastruktur wird an der Universität Wien weiterentwickelt [Abt Jf]. Daneben nimmt die Universitätsbibliothek auch an den Infrastruktur-Projekten OpenAIRE und e-Infrastructures Austria teil [Abt Ja]. Weiterhin ist die Abteilungsleitung die Kontaktperson für das österreichische "Netzwerk für RepositorienmanagerInnen"⁸² [BF J]. Die Mitarbeiterx führen eine Liste ihrer Publikationen⁸³.

Technische Basis von Phaidra ist Fedora, dieses wird auf Apache- und Tomcat-Servern unter Linux betrieben [Gan17, S. 110], die Suche basiert auf Apache Solr [Abt Jg]. Eine Zusammen-

⁷⁶<http://datamanagement.univie.ac.at/ueber-phaidra/technik/>

⁷⁷<http://datamanagement.univie.ac.at/ueber-phaidra/zugang-zu-phaidra/>

⁷⁸<http://datamanagement.univie.ac.at/>

⁷⁹<https://datamanagement.univie.ac.at/rdm/datenmanagementplaene/>

⁸⁰<http://datamanagement.univie.ac.at/>

⁸¹<http://datamanagement.univie.ac.at/ueber-uns/>

⁸²Mehr Informationen zum Netzwerk Repmannet: <https://ubifo.at/images/repmannet.pdf>

⁸³<http://datamanagement.univie.ac.at/ueber-phaidra/publikationen/>

fassung ist auf der Webseite⁸⁴ verfügbar und die technische Dokumentation von Phaidra ist in github⁸⁵ hinterlegt. Alle verwendeten Elemente sind typischerweise offen im Sinne der Open Source Bewegung, explizite Aussagen werden dazu jedoch nicht getroffen.

K6 Gebühren und Finanzierung

Die Abgabe von Daten ist kostenlos ([Abt Jf] und [Ami J]), während der Erstellung der Analyse wurden keine Kosten für den Download von Daten verlangt. Zur Finanzierung von Phaidra gibt es weder Hinweise auf der Webseite noch in den Jahresberichten der Universitätsbibliothek⁸⁶.

5.3. Zenodo

K1 Veröffentlichung und Nachnutzung

Zenodo akzeptiert alle Dateiformate und bezieht auch explizit solche ein, die "preservation unfriendly" [Zen Jd] sind, das Größenlimit pro Datensatz liegt bei 50GB. Der Upload erfolgt per Formular (siehe Screenshot Zen008_Upload) oder über eine REST Schnittstelle [Zen Jh]. Das Metadatenschema ist DataCite-konform [Zen Jg] und wird als JSON gespeichert [Zen Je]. Die gekennzeichneten Pflichtangaben entsprechen den "mandatory terms" der DataCite Richtlinie, optionale MD setzen sich zusammen aus den empfohlenen Angaben von DataCite und lokalen Ergänzungen [Zen Jg]. Für die Vergabe von Schlagworten ist keine Systematik hinterlegt (siehe Screenshot Zen009_Schlagworte), kontrollierte Vokabulare werden aber für Felder wie Geldgeber (FundRef) und Stipendien (OpenAIRE) angeboten [Zen Jg].

Es besteht keine Verpflichtung zur Dokumentation der FD, die Möglichkeiten zum Upload von zugehöriger Software werden aber gegeben [Zen Ja].

Die Verlinkung auf zugehörige Datenpublikationen wird im Eingabeformular im Reiter "Related/alternate identifiers" ermöglicht [Zen Jf]. Für die Verlinkung werden diverse Identifikatoren akzeptiert, darunter DOI und URN (ebd.). Der Screenshot Zen003_VerwandtePublikationen zeigt die Möglichkeiten die Beziehung zwischen FD und Publikationen zu definieren.

Zenodo unterstützt die Data Citation Principles⁸⁷ und zeigt einen Zitationsvorschlag auf Datensebene an (siehe Screenshot Zen011_Zitationsvorschlag). Als PID werden DOI verwendet. Existiert keine externe DOI vergibt Zenodo automatisch eine neue DOI, da es sich um eine Pflichtangabe in den Metadaten handelt [Zen Jg].

Eine weitere Pflichtangabe ist die Lizenzbedingung. Wenn keine offene Lizenzen benutzt wird, kann eine restriktivere Lizenz im "Description" Feld durch Nutzerx spezifiziert werden [Zen Jd]. Offene Lizenzen wie Creative Commons werden empfohlen (siehe Screenshot Zen012_Lizenzen), Zenodo bietet für die breite Nutzerxbasis aber verschiedenste Zugriffsbeschränkungen und Lizenzen an [Zen Jb].

Metadaten sind explizit CC0 [Zen Ja] bzw. Public Domain [Zen Jg] lizenziert.

Verlinkungen zu Publikationen des Forscherx und Publikationen, die die FD nachnutzen sind wie oben beschrieben über das Metadatum "Specify identifiers of related publications and datasets" [Zen Jf] möglich (siehe Screenshot Zen003_VerwandtePublikationen). Für die Verknüpfung verwandter Projekte bietet Zenodo die Option eigene Collections oder Communities anzulegen [Zen Jb], siehe dazu auch Screenshot Zen002_Communities.

Die Anerkennung der wissenschaftlichen Leistung erfolgt vor allem über Anreize. So werden z.B. nur FD unter offenen Lizenzen auf der Front-Seite präsentiert [Zen Jc]. Weiterhin wird eine vereinfachte Rechtfertigung gegenüber Mittelgebern genannt: "OpenAIRE integrates [research

⁸⁴<http://datamanagement.univie.ac.at/ueber-phaidra/technik/>

⁸⁵<https://github.com/phaidra>

⁸⁶http://bibliothek.univie.ac.at/zahlen_und_daten.html

⁸⁷<https://www.force11.org/datacitationprinciples>

results] into existing reporting lines to funding agencies like the European Commission." [Zen Ja]. Daneben ist die direkte Verknüpfung mit dem Orcid Konto möglich, für Neu-Registrierende Forscherx kann darüber sogar der Login zu Zenodo ersetzt werden (siehe Screenshot Zen013_Orcid). Die ORCID kann in den Autorenfeldern eingegeben werden wenn FD hochgeladen werden [Now17]. Statt Zugriffs- oder Downloadzahlen werden auf Datensatzeben Altmetrics angezeigt (siehe Screenshot Zen004_AltMetric).

K2 Zugriff und Auffindbarkeit

FD können als Geschlossen ("closed"), Öffentlich ("open") oder unter Embargo hochgeladen werden [Zen Jd]. Den Zugang für "closed" Datensätze verwaltet die hochladende Person (ebd.). Die Datei wird jedoch nicht verschlüsselt, daher wird vom Upload von vertraulichen Daten abgeraten [Zen Je]. Die Nachnutzung der Daten ist nur im zivilen Bereich erlaubt: "Access to Zenodo's content is open to all, for non-military purposes only." [Zen Ji].

Die Zielgruppe sind alle Forscherx, die kein eigenes FDR zur Verfügung haben, sei es institutionell oder disziplinar [Zen Jc]. Entsprechend ist jede Person zum Upload berechtigt, dafür ist nur eine kostenlose Registrierung nötig [Zen Jd].

Um Datensätze unter Embargo zu stellen muss die hochladende Person ein Enddatum angeben, mit dessen Ablauf die Datei automatisch veröffentlicht wird [Zen Jd]. Alternativ kann das Embargo durch die Publikation eines verlinkten Artikels aufgehoben werden [Zen Ja]. Während der Frist sind auf der Landingpage die vollen Metadaten und die Dauer des Embargos angegeben (siehe Screenshot Zen014_Embargo).

Für die Suche nach FD steht eine einfache Schlitzsuche (siehe Screenshot Zen005_Suchschlitz) zur Verfügung. Zur Suche wird keine Hilfeseite angeboten, in der Analyse konnten keine aussagekräftigen Ergebnisse bei Tests mit Trunkierung oder Wildcards erreicht werden. Die attributspezifische Suche ist über die integrierte Suchmaschine Elasticsearch möglich [Zen Je]. Eine Facettierung ist u.a. verfügbar für Zugangsbedingungen und Dateitypen (siehe Screenshot Zen006_Facettierung). Die Landingpages werden nicht explizit als indexierbar bezeichnet, aber eine Known-Item-Search in Google war erfolgreich (siehe Screenshot Zen016_IndexingGoogle). Die Zenodo Metadaten sind bereit für die Integration in Zitationsdatenbanken und werden über DataCite gemeldet [Zen Jg]. Über die OAI-PMH Schnittstelle können Datensätze und Sammlungen geharvested werden (ebd.), dies gilt für alle Metadatenfelder [Zen Jd].

Die REST API ermöglicht ebenfalls die Abfrage von Metadaten [Zen Jg] und kann wie erwähnt für den Upload von FD verwendet werden. Noch in Planung befindlich sind REST-basierte Suchfunktionen für FD, Communities usw. [Zen Jh].

K3 Transparenz und Qualitätssicherung

Zenodo bietet ihren Nutzerx Terms of Use⁸⁸ und Policies⁸⁹. Weitere relevante Details sind über eine Anzahl von Hilfeseiten verteilt. Anstelle einer Leistungsvereinbarung hält sich Zenodo an die FAIR Principles [Zen Jg].

Daten werden vor und im Ingest nicht durch das FDR validiert, stattdessen erfolgt eine Bereitstellung "as-is" [Zen Jd]. In den Terms of Use werden Datengeberx als Verantwortliche für die Datenqualität benannt [Zen Ji].

Die Versionierung von Datensätzen ist möglich, sowohl als Angabe in den Metadaten [loa17] (z.B. für Softwareversionen) als auch für die Inhalte der Uploads an sich [Zen Jd]. Die Verknüpfung erfolgt über verschiedene Identifier: Eine allgemeine DOI, die für alle Versionen eines Datensatzes gilt und eine weiteren DOI für jede neue Version [Zen Jc]. Für Nutzerx wird die Vergabe der DOI im Kapitel "DOI Versioning" in den FAQs verdeutlicht.

Die Löschung von Daten ist nur in Ausnahmefällen vorgesehen und erfolgt nur nach Angabe

⁸⁸<http://about.zenodo.org/terms/>

⁸⁹<http://about.zenodo.org/policies/>

eines Grundes durch Datengeberx. Die DOI und URL des ursprünglichen Datensatzes bleiben bestehen und zeigen danach auf eine Tombstone page, also eine leere Landingpage, die keinen Zugriff auf die Daten erlaubt. Der Grund der Löschung wird dort vermerkt [Zen Jd].

K3.5 Sicherstellung langfristiger Zugänglichkeit

Der Abschnitt "longevity" in den General Policies⁹⁰ versichert die Sicherstellung langfristiger Zugänglichkeit gegenüber den Nutzerx. Einige der Sicherheitsmaßnahmen sind ortsverteilte Speicherung (in Genf und Budapest), tägliche Backups auf Bandspeicher und regelmäßige Checksum Prüfungen [Zen Jd]. Die Sicherungsdauer wird am Träger CERN festgemacht und auf mindestens 20 Jahre festgelegt [Zen Jd]. Für den Fall einer nötigen Nachfolge wird die Weitergabe in andere institutionelle oder disziplinäre Repositorien garantiert [Zen Jd].

Zenodo ist nicht zertifiziert.

K4 Nutzererlebnis und Support

Die Usability des FDR ist durch die fehlende Erklärung der Suchfunktionen eingeschränkt. Fehlermeldungen die aus Fehleingaben entstehen sind nicht sprechend gestaltet, siehe Zen015_Usability-Error. Barrierefreiheit ist nicht gegeben. Die Browsererweiterung aXe findet 173 Verstöße gegen die Barrierefreiheit in der Anzeige der Suchergebnisse (siehe Screenshot Zen017_Accessibility-Suche), die meisten für fehlenden Farbkontrast und fehlende Beschriftung in formtags. Weitere 46 Verstöße werden auf Datensatzebene (siehe Screenshot Zen018_Accessibility-Datensatz) gefunden, auch hier vor allem fehlender Farbkontrast, sowie fehlende Textalternativen und Linkbeschriftungen. Die manuelle Betrachtung des FDR ergab, dass die inhaltliche Beschreibung von Videos lediglich über Freitext möglich ist. Eine Stichprobenprüfung für jeweils 3 Datensätze aus Type Image (Photo), Video/Audio fand keine Textalternativen. Die Suche ist nicht mit Tastatur allein bedienbar und bietet weder Eingabehilfe noch Fehlerkorrektur.

Eine Vorschau ist nur für einzelne Bilder möglich (siehe Screenshot Zen019_PreviewBild), alle anderen Arten von Ressourcen müssen ungesehen heruntergeladen werden.

Als Ressourcen für die Hilfe bietet Zenodo die Frequently Asked Questions⁹¹. Supportanfragen können über ein Kontaktformular⁹² gesendet werden.

Es findet keine Beratung zum Forschungsdatenmanagement statt.

K5 Organisation und Infrastruktur

Zenodo wird durch die IT-Abteilung von CERN betreut, spezifisch durch die "Collaboration Devices and Applications Group" und deren "Digital Repositories Section". Nach eigener Aussage hat CERN "some of the world's top experts in running large scale research data infrastructures and digital repositories" [Zen Je] angestellt. Die Webseite der Abteilung enthält eine Liste der aktuellen Mitarbeiter⁹³.

Aussagen zur Langfristigkeit von Zenodo werden direkt an CERN gekoppelt: "Zenodo is hosted by CERN which has existed since 1954 and currently has an experimental programme defined for the next 20+ years." [Zen Je].

Eine volle Liste der Geldgeber und Träger des FDR ist auf den Webseiten veröffentlicht und beinhaltet die Europäische Kommission über diverse Projekte und CERN selbst [Zen Je].

Zenodo beteiligt sich aktiv an diversen EU-geförderten Projekten, darunter OpenAIRE und EU-DAT [Zen Je]. Zudem ist es Unterzeichner der Joint Declaration of Data Citation Principles⁹⁴. Die Trägerorganisation CERN listet zusätzlich diverse Mitgliedschaften, darunter ORCID, die Research Data Alliance und SCOAP3 [Zen Je].

⁹⁰<http://about.zenodo.org/policies/>

⁹¹<http://help.zenodo.org/>

⁹²<https://zenodo.org/support>

⁹³<https://phonebook.cern.ch/phonebook/#search/?query=IT-CDA-DR>

⁹⁴<https://www.force11.org/datacitation/endorsements>

Zenodos technische Infrastruktur basiert auf Invenio, welches wiederum auf Python und Flask basiert [Zen Je]. Invenio wird als "battle-tested repository software"[Zen Jb] beschrieben, mit der auch die Daten aus dem Large Hadron Collider verwaltet werden.

Im September 2017 Update wurde die Geschwindigkeit der Suche optimiert indem Elasticsearch eingebunden wurde [Zen17]. Das Produktiv- und Testsystem werden getrennt gehalten um neue Features vor Release zu testen ohne den Zenodo Service zu unterbrechen [Zen Je]. Zum Back-Up der FD werden jede Nacht Kopien angefertigt, Metadaten und PID werden alle 12 Stunden gesichert und wöchentlich auf Band geschrieben [Zen Je]. Die Überwachung der Server-Hardware geschieht durch die Software die auch CERN nutzt, darunter Apache Flume und Hadoop [Zen Je].

Der aktuelle Status von Zenodo kann Online eingesehen werden⁹⁵, ausgelesen werden Datei Up- und Downloads, Suchen, Server Uptime usw. Für den Stand vom 30.12.2017 siehe Screenshot Zen020_Status.

Die technische Dokumentation umfasst neben den APIs⁹⁶ auch den vollständigen Zenodo SourceCode auf github⁹⁷. Zenodo wird beschrieben als "[...] fully run on open source products all the way through." [Zen Je].

K6 Gebühren und Finanzierung

Derzeit werden keine Gebühren für die Nutzung von Zenodo verlangt [Zen Ji]. Im Kontext des Uploads von Datensätzen über 50GB wird jedoch erwähnt: "Please be aware that we cannot offer infinite space for free, so donations towards sustainability are encouraged." [Zen Jc].

Abgesehen von der Versicherung, dass Zenodo so lange wie CERN bestehen wird [Zen Jd] werden keine Aussagen zum langfristigen Konzept gemacht. Neben den offengelegten Trägern und Geldgebern [Zen Je] werden Spenden zum Weiterbetrieb von Zenodo über die Spendenseite "Giving to CERN"⁹⁸ akzeptiert.

5.4. Brown Digital Repository

K1 Veröffentlichung und Nachnutzung

Das BDR ist kein reines FDR, es akzeptiert Publikationen genauso wie Video und Digitalisate [Bro Jf]. Akzeptiert werden Ressourcen in allen Dateiformaten, von denen eine Anzahl vom System erkannt und direkt im Web bereitgestellt werden kann, darunter .pdf und die Videoformate .flv, .mov und .mp4. Die Größe der Datensätze sollte unter 4GB bleiben, Ausnahmen sind nach Absprache möglich. [Bro Jf].

Als Möglichkeit für den Upload wird ein "Upload Screen" [Bro Je] in den FAQ erwähnt, weitere Informationen gibt es für Externe Benutzerx nicht.

Die Integrationsprüfung wird durch Bibliothekarx durchgeführt, gleichzeitig wird u.a. auch nach Sicherheitsbedrohungen in den Dateien ermittelt [Bro Jf].

Auf den Hilfeseiten der Bibliotheksabteilung "Digital Production Services" werden Descriptive, Administrative, Technical, Structural, Preservation und Rights Metadaten unterschieden [Bro Ji]. Weiterhin gibt das BDR an, als Metadatenschema eine Kombination von Dublin Core, MODS und METS zu verwenden [Bro Jf].

Die Verlinkung auf eine zugehörige Datenpublikation, Publikation der Ergebnisse oder nachnutzende Publikation ist über das Metadatum "Related Item" möglich, siehe dazu den Screenshot

⁹⁵<https://status.zenodo.org/>

⁹⁶<http://developers.zenodo.org>

⁹⁷<https://github.com/zenodo/zenodo>

⁹⁸<https://giving.web.cern.ch/civicrm/contribute/transact?reset=1&id=20>

BDR004_Verlinkung oder die BDR Webseite auf Datensatzebene unter Full Metadata⁹⁹. Die Verknüpfung zugehöriger Projekte geschieht über Sammlungen (siehe Screenshot BDR006_Sammlungen).

Ebenfalls auf Datensatzebene wird ein Zitationsvorschlag angeboten (siehe Screenshot BDR003_Zitationsvorschlag). Die Zitierfähigkeit wird dabei durch DOI gewährleistet, die von der Bibliothek selbst vergeben werden [Bro Jd].

In der Policy des BDR wird nur das Copyright der Autoren vermerkt [Bro Jf], in einzelnen Datensätzen wird aber deutlich, dass auch offene Lizenzen wie CC und GNU-GPL (siehe Screenshots BDR010_Lizenz-CC und BDR011_Lizenz-GNU) vergeben werden können.

Als Anreiz die FD zu veröffentlichen werden die Vorteile einer DOI genannt, welche Datensätze zitierfähig und sichtbar machen und außerdem in Impact Messungen registriert werden [Bro Jd]. Es gibt keine Hinweise auf eine Integration der ORCID im BDR, obwohl die Brown University Institutionelles Mitglied ist und für Angehörige ORCID zentral vergeben kann [Rai15]. Ebenso werden Zitationsraten und sonstige Metriken nur über den Umweg DOI-Vergabe angeboten: Die FD können dank DOI in der Berechnung von Impact- Kennzahlen integriert werden, als Beispiele werden ImpactStory und Google Scholar genannt [Bro Jd]. Laut der Strategischen Planung für 2017/18 ist ein Pilotprojekt für die Implementation einer Nutzungsstatistik angedacht [Bro Jj, S. 3].

K2 Zugriff und Auffindbarkeit

Der Zugang kann durch Nutzerx auf drei Ebenen geregelt werden: Öffentlich, nur Universitätsangehörige und Privat. Eine Änderung der Zugriffsform ist jederzeit möglich [Bro Je]. Weitere Gruppen können auf Anfrage eingerichtet werden [Bro Jf]. Zum Upload sind alle Universitätsangehörigen berechtigt, Studierende jedoch nur wenn die Arbeit im Rahmen von Praxisrelevanten Projekten erstellt wird [Bro Jf]. Die Zugriffssteuerung geschieht über Shibboleth (ebd.).

Die ebenfalls im BDR gespeicherten Dissertationen nutzen den Status "Brown community" als Embargofunktion, dabei bleiben die Metadaten öffentlich [Cai17b]. Zu Embargo für FD wird keine Aussage getroffen.

Die Suche nach Datensätzen ist über eine Einfache Suche über alle Felder oder die Angaben in Title, Author und Subject möglich (siehe Screenshot BDR007_Suche). Eine Sortierung der Treffer ist nicht möglich. Alternativ kann die Search API mit einem Solr query abgefragt werden [Bro Jc]. Das BDR bietet keine Suchhilfe. Tests mit Wildcards und Trunkierung verliefen ohne Fehlermeldung (siehe Screenshot BDR012_WildcardTrunkierung), ein Test mit exakten Suchbegriffen in Anführungszeichen war erfolgreich (siehe Screenshot BDR013_ExactSearch). Eine Einschränkung der Ergebnisse ist über die Felder Format, Keywords und Genre möglich (siehe Screenshot BDR008_Facetten). Die Landingpages sind indexierbar sofern der Zugriff auf Public gesetzt wird [Bro Je]. Außerdem erfolgt ein Austausch von MD mit Suchmaschinen und anderen Repositorien [Bro Jf]. Laut der Strategischen Planung für 2017/18 ist ein Pilotprojekt für Volltextsuche vorgesehen [Bro Jj, S. 3].

Insgesamt bietet das BDR drei APIs: Die Item API fragt Metadaten der Datensätze ab, die Collection API ruft Sammlungen ab und die oben erwähnte Search API funktioniert analog der Suchfunktion auf der Webseite [Bro Jc].

K3 Transparenz und Qualitätssicherung

Das BDR hat eine explizite Policy¹⁰⁰ und verweist auf die allgemeinen Richtlinien "Copyright and Fair Use at Brown"¹⁰¹ und die "Acceptable Use Policy"¹⁰² der IT-Abteilung.

⁹⁹ Beispieldatensatz: <https://doi.org/10.7301/Z0668B3H>

¹⁰⁰ <https://repository.library.brown.edu/studio/policies>

¹⁰¹ <https://www.brown.edu/about/administration/copyright>

¹⁰² <https://it.brown.edu/computing-policies/acceptable-use-policy>

Alle eingespeisten Daten werden wie erwähnt durch Bibliothekarx validiert [Bro Jf]. Zur Versionierung wird nur im Kontext der LZA informiert, dort wird bei jeglicher Änderungen an den Daten eine neue Version und ein Änderungsprotokoll erstellt [Bro Jf]. Auf eine mögliche Versionierung durch Nutzerx deutet die Vergabe der PID hin: "New versions of existing objects will use the same permanent links unless an entirely new object is created." [Bro Jf].

Jeder Nutzerx kann die Löschung seiner Datensätze mit Beendigung der Universitätsangehörigkeit anfordern, diesem wird Folge getragen sofern keine anderen Rechte oder Verträge verletzt werden [Bro Jf].

Im Rahmen der LZA werden ortsverteilte Backups gepflegt und auf Datenebene regelmäßig Migration durchgeführt [Bro Jb]. Ist im Rahmen der Migration eine Änderung am Bitstream nötig, werden Datengeberx informiert und haben die Option die Daten stattdessen aus dem BDR zu entfernen [Bro Jf].

Das Brown Digital Repository ist nicht zertifiziert.

K4 Nutzererlebnis und Support

Das BDR bietet keine Suchhilfe, nutzt aber etablierte Suchoperatoren und Darstellungsformen. Die Webseite ist nicht barrierefrei. Die Browsererweiterung aXe findet 44 Verstöße gegen die Barrierefreiheit in der Anzeige der Suchergebnisse (siehe Screenshot BDR014_Accessibility-Suche), Probleme sind vor allem fehlender Farbkontrast und fehlende Beschriftung von Links. Weitere 3 Verstöße werden auf Datensatzebene gefunden (siehe Screenshot BDR015_Accessibility-Datensatz), auch hier vor allem fehlender Farbkontrast. Weiterhin sind die betrachteten Videos nicht inhaltlich beschrieben. Die Universität hat einen Closed Captioning Service¹⁰³ und die zuständige Abteilung informiert, dass alle Online veröffentlichten Videos Untertitel beinhalten sollten [Bro Ja]. Im Gegensatz dazu bietet das BDR keine Textalternativen (Stichprobenprüfung für jeweils 3 Datensätze aus Format Still image, Moving image und Sound recording). Die Anpassung von Textgröße oder Farbschema ist nicht möglich. Die Navigation der Webseite ist mit Tastatur allein möglich. In den strategischen Zielen für 2017/18 werden eine Selbstprüfung anhand der US Accessibility Standards¹⁰⁴ und "accessibility checks" für alle Projekte angekündigt [Bro Jj].

Die Bibliotheksabteilung CDS hat einen "Data Visualization Coordinator" zur Unterstützung von Forscherx, die ihre Daten visualisieren möchten [Bro Jg]. Für Videos steht eine Streaminoption zu Verfügung [Bro Je], siehe auch Screenshot BDR005_Vorschau.

Als Hilfeseiten stehen Frequently Asked Questions¹⁰⁵, Informationen zu DOI¹⁰⁶ und APIs¹⁰⁷ bereit, zusätzlich werden vor Ort Workshops zu diversen Themen angeboten, darunter auch Datenvisualisierung¹⁰⁸. Für Supportanfragen verweist die Verlinkung "Contact the Digital Repository" im Footer der Webseite auf die e-Mail Adresse der zuständigen Abteilung.

Eine Beratung zum FDM erfolgt über das Center for Digital Scholarship zu Themen wie Data Curation, Metadata Creation oder Digital Humanities allgemein [Bro Jg] oder über die Digital Production Services u.a. zu den Themen Digitalisierung, inkl. Metadata consulting oder Metadata application [Bro Ji]. Als aktive Hilfestellung wird Unterstützung beim Einspeisen von Daten in das BDR angeboten [Bro Jk], sowie diverse "Data Management and Sharing Services", darunter z.B. Vorlagen für Datenmanagementpläne¹⁰⁹. Bibliotheks-extern bieten die Computing & Information Services (Brown IT Abteilung) u.a. Dienstleistungen zur Aufnahme

¹⁰³<https://it.brown.edu/services/type/closed-captioning-digital-content>

¹⁰⁴Sog. "Voluntary Product Accessibility Templates": <https://www.section508.gov/content/sell/vpat>

¹⁰⁵<https://repository.library.brown.edu/studio/faq>

¹⁰⁶<https://repository.library.brown.edu/studio/doi/>

¹⁰⁷<https://repository.library.brown.edu/studio/api-docs/>

¹⁰⁸<https://library.brown.edu/cds/activities/workshops/#data-analysis>

¹⁰⁹https://library.brown.edu/info/data_management_plans

und Bearbeitung von Videos¹¹⁰.

Im Bereich der aktuellen Richtlinien und Standards verweist das BDR u.a. auf "Resources for Data Management Best Practices"¹¹¹, Data Curation Best Practices¹¹² und re3data¹¹³.

K5 Organisation und Infrastruktur

Das Ziel des BDR ist "[...] to gather, index, store, preserve, and make available digital assets produced via the scholarly, instructional, research, and administrative activities at Brown." [Bro Jb], die direkte Anbindung an die Bibliothek ist dafür angemessen. Der zuständige Träger wird auf der Webseite deutlich genannt [Bro Jb]. Die fachliche Zuständigkeit wird nur für die Digital Technologies Abteilung explizit erwähnt [Bro JI], und ist im Falle des Center for Digital Scholarship implizit über die Mitarbeiter und angebotenen Dienstleistungen.

Auf der Personalseite der Bibliothek¹¹⁴ wird das BDR nicht aufgeführt. Die Abteilung Center for Digital Scholarship hat 8 Mitarbeiter¹¹⁵ inkl. einem Digital Repository Manager und einem Data Visualization Coordinator. Die Abteilung Digital Technologies hat 3 Mitarbeiter für das BDR¹¹⁶: Je ein Digital Repository Manager, Digital Library Programmer und Digital Preservation Librarian.

Die Weiterentwicklung der Infrastruktur ist das Kernstück der Abteilung CDS, zusammen mit der Entwicklung von Best Practices für LZA und FDM [Bro Jh]. Ebendiese Abteilung führt auch diverse Digitalisierungsprojekte in Kooperation mit Angehörigen der Fachdisziplinen durch¹¹⁷. Weiterhin arbeiten CDS Mitarbeiter an der Entwicklung von FDM Standards mit, die Mitarbeiter sitzen in entsprechenden Gremien und Testen neue Standards [Bro Jh].

Das BDR basiert mit Stand September 2017 auf Fedora 3, dies ist aber veraltet und soll aber bald ersetzt werden [Cai17a], und nutzt Loris als IIF-konformen¹¹⁸ Server für Bilddaten [Cai16b]. Das Server-Betriebssystem ist Red Hat Enterprise Linux in der Version 7, dabei werden Loris (Bilder), Solr (Suche) und Fedora (Repositorium) auf verschiedenen Servern getrennt betrieben, welches die Stabilität und Sicherheit des Systems erhöht [Cai17c]. Die Abteilung die das BDR technisch betreut hat "Digital Technologies Strategic Goals, 2017-18" aufgestellt, die auch das BDR betreffen. Geplant sind z.B. ein Pilotprojekt für Volltextsuche und die Implementation einer Nutzungsstatistik [Bro Jj, S. 3].

K6 Gebühren und Finanzierung

Zu Gebühren werden auf der Webseite des BDR keine Aussagen getroffen.

Als langfristiges technisches Konzept lag bisher der "Brown University Libraries Technology Plan, 2015-2017"¹¹⁹ vor, der auch Pläne für das BDR enthielt. Weitere Hinweise sind für externe Nutzer nicht vorhanden.

5.5. LSE Research Online

K1 Veröffentlichung und Nachnutzung

Das LSERO steht allen Mitarbeitern und Doktoranden der Hochschule zum Upload zur Verfügung [LSE Je]. Akzeptiert werden Publikationen und Forschungsdaten, sowie graue Literatur

¹¹⁰ <https://it.brown.edu/services/type/video-recording-editing-and-delivery>

¹¹¹ https://library.brown.edu/info/data_management_resources

¹¹² https://library.brown.edu/info/data_management_curation

¹¹³ https://library.brown.edu/info/data_management_depositing

¹¹⁴ <https://library.brown.edu/about/deptlist.php#dept12>

¹¹⁵ <https://library.brown.edu/cds/about/people/>

¹¹⁶ <https://library.brown.edu/dt/>

¹¹⁷ <https://library.brown.edu/cds/projects/>

¹¹⁸ Mehr zum International Image Interoperability Framework: <http://iiif.io/about/>

¹¹⁹ <https://library.brown.edu/dt/LORITechnologyPlan2015.pdf>

[LSE14], aber auch Blogbeiträge und Zeitungsartikel [LSE Je]. Die Bibliothek hat eine Liste mit empfohlenen offenen Formaten zusammengestellt¹²⁰, darunter WebM für audiovisuelle FD, zu den zulässigen Formaten für die Einspeisung in LSERO wird jedoch keine Aussage getroffen. Nutzer können ihre FD auf unspezifiziertem Weg selbst hochladen oder die Dateien per e-Mail einsenden [LSE Je]. Die Daten-Integritätsprüfung wird von den Mitarbeiterx übernommen [LSE Je].

Laut re3data wird im LSERO Dublin Core verwendet [re317b], dies wird auf der Webseite jedoch nicht deutlich gemacht. Einen Hinweis auf ergänzende Metadaten die erfasst werden liefern Ayre und Madjarevic auf einem Poster: Dort werden u.a. Metadaten zu Drittmittelgebern genannt [AM14]. Die Metadaten, die von Bibliothekarx erstellt werden, beinhalten mindestens die Projektdetails die von der RCUK Richtlinie gefordert werden und weitere Angaben wie Lizenzen [Wil17, S. 3].

Unterstützung beim Einspeisen von Blog Posts erhalten die Mitarbeiterx des LSERO über automatisch ausgelesene und vorausgefüllte Metadaten, dies ist aber noch in Entwicklung [Wil17, S. 4].

Die erweiterte Suche zeigt, dass eine Systematik für das Feld Subject hinterlegt ist (siehe Screenshot LSE007_AdvancedSearch), weitere Details werden nicht genannt.

Zur Zitation von Datensätzen bietet das Repositorium einen Zitationsvorschlag auf Datensatzebene (siehe Screenshot LSE001_AnzeigeDatensatz).

DOI können durch die Bibliothek vergeben werden, wenn die Objekte in LSERO gespeichert sind, ein Bezug zur Hochschule besteht, keine andere Stelle eine DOI vergibt und keine weiteren Änderungen am Inhalt erwartet werden [LSE Ja]. Auf Datensatzebene wird für jeden Eintrag eine URI angezeigt (siehe Screenshot LSE001_AnzeigeDatensatz).

Die Nutzungsbedingungen werden in den FAQ festgelegt, Nutzerx werden informiert das das Copyright auch für Datensätze im Internet zutrifft und durch diese Art der Veröffentlichung keine Sonderrechte entstehen [LSE Je].

Williams erwähnt das eine "preferred Creative Commons attribution" [Wil17, S. 4] in den Metadaten angegeben werden kann, weitere Aussagen zu offenen Lizenzen werden nicht getroffen. Autoren werden dazu aufgefordert den Link zur Veröffentlichung der zugehörigen Publikationen anzugeben [LSE Je]. Eine Verknüpfung zugehöriger Projekte ist über Sammlungen möglich (siehe Screenshot LSE001_AnzeigeDatensatz).

LSERO zeigt die Downloadzahlen seit Upload des Datensatzes an (siehe Screenshot LSE002_Statistik) und bezeichnet dies als "[...] useful data alongside traditional citation counts." [Ayr13].

Als Anreize für Forscherx werden vor allem die erhöhte Zugänglichkeit der Forschungsdaten angeführt: "Depositing research in LSERO increases its visibility and the potential for increased citations; it also creates opportunities for collaboration with other researchers." [LSE14]. Außerdem werden in LSERO gepflegte Publikationen automatisch auf den Mitarbeiterseiten, den sog. Experts profiles der Hochschule angezeigt [LSE Je].

Es ist möglich LSERO Datensätze über BibTex in ORCID zu übernehmen [LSE Jc] und "Adopting a reliable author name authority eg, ORCID" [AM14] wird auf einem Poster zum LSERO empfohlen, aber eine Suchmöglichkeit oder Anzeige der ORCID konnte nicht gefunden werden. Altmetrics werden auf Datensatzebenen angezeigt (siehe Screenshot LSE008_Altmetric) wenn eine DOI vorhanden ist, die Kennzahl wird als eine "score of attention" [Ayr13] beschrieben.

K2 Zugriff und Auffindbarkeit

Zu den Zugangformen wird keine Aussage getroffen, laut re3data sind aber die Optionen offen und eingeschränkt möglich [re317b]. Ein Blogartikel erwähnt ebenfalls die Möglichkeit

¹²⁰<http://www.lse.ac.uk/library/research-support/research-data-management/file-formats>

"restricted-access items" [Fla17] hochzuladen. Flanagan erwähnt ebenfalls, dass eine Embargofunktion existiert, die mindestens zwei Jahre Sperrfrist erlaubt (ebd.).

Für die Recherche stehen neben einer eingeschränkten Browse-Funktion (siehe Screenshot LSE003_Browse) eine einfache und eine erweiterte Suche zu Verfügung. Letztere bietet Auswahlmöglichkeiten über eine Freitextsuche in Feldern wie Documents, Title und Abstract, sowie Einschränkung auf Subjects, Item Type, Status usw. (siehe Screenshot LSE007_AdvancedSearch). Zu keinem der Einstiege wird eine Suchhilfe angeboten. Die Trunkierung scheint möglich zu sein (siehe Screenshot LSE009_Trunkierung). Es gibt keine expliziten Aussage auf der Webseite aber das Suchfeld "Documents" impliziert die Möglichkeit einer inhaltliche Suche in Textdokumenten.

Die Landingpage wird mindestens von Google und Google Scholar indexiert ([LSE Je] und [LF16]).

Das Repositorium hat eine OAI-PMH Schnittstelle [LSE Jg], die gemeinsam mit der base URL im Footer verlinkt ist und eine Abfrage der Metadaten erlaubt [LSE Je].

K3 Transparenz und Qualitätssicherung

Das Mission Statement des LSERO ist in den FAQ enthalten¹²¹, zusätzlich gibt es eine Take-down Policy¹²². Die Hochschule hat außerdem eine LSE Open Access Publications Policy¹²³ und eine Research Data Management Policy¹²⁴, sowie allgemeine Nutzungsbedingungen für die Webseiten der Universität¹²⁵.

Aussagen zur Löschung von Datensätzen werden nur in den OAI PMH Daten gemacht, dort heisst es: "Deleted Record Policy: persistent" [LSE Jg]. Weitere Informationen zur Löschung werden davon abgesehen nur bei Verstößen gegen das Copyright genannt [LSE Jf].

Auf eine Möglichkeit der Löschung weist eine Aussage in den FAQ hin: "A citation/citations to the Work(s) will always remain visible in the repository during its lifetime." [LSE Je]. Dies kann aber nicht durch weitere Aussagen bestätigt werden.

Zur Sicherstellung der langfristigen Zugänglichkeit wird erwähnt, dass Datensätze übersetzt oder inhaltlich umgestellt werden können um Lesbarkeit zu garantieren [LSE Je].

Das LSERO ist nicht zertifiziert.

K4 Nutzererlebnis und Support

Zu Erhöhung der Usability können die Metadaten der Datensätze im BibTex Format heruntergeladen werden und somit in ein vorhandenes ORCID Profil integriert werden [LSE Jc]. Alle Einträge werden im Discovery System der Hochschule angezeigt und automatisch auf den Hochschul-Webseiten der Mitarbeiter eingepflegt [AM14].

Der Aspekt der intuitiven Bedienung ist z.B. durch die fehlende Möglichkeit eine leeren Suche um alle Suchergebnisse anzeigen zu lassen eingeschränkt (siehe Screenshot LSE010_LeereSuche) zudem ist eine Anpassung der Suchterme auf der Ergebnisseite nicht möglich (siehe Screenshot LSE004_AnzahlErgebnisse). Die Darstellungsformen sind etabliert, entsprechen aber nicht den modernen Ausführungen, ein Beispiel dafür ist das Lupensymbol für "View Item" (siehe Screenshot LSE001_AnzeigeDatensatz). Die Lernkurve ist eher hoch, u.a. da Tooltips nicht immer Aussagekräftig sind (siehe Screenshot LSE011_Tooltips) und wegen der fehlenden Suchhilfe.

Die Hochschule hat einen Accessibility Absatz in den Terms of Use [Lon J], dennoch ergab die

¹²¹<http://eprints.lse.ac.uk/faq.html>

¹²²<http://eprints.lse.ac.uk/takedown.html>

¹²³https://info.lse.ac.uk/staff/services/Policies-and-procedures/Assets/Documents/opeAccPub.pdf?from_serp=1

¹²⁴<https://info.lse.ac.uk/staff/services/Policies-and-procedures/Assets/Documents/resDatManPol.pdf>

¹²⁵<http://www.lse.ac.uk/lse-information/terms-of-use>

manuelle Betrachtung, dass es keine Möglichkeit zu inhaltlichen Beschreibung von Videos und eine Funktion für das Einstellen von Transkripten oder Untertitel zu geben scheint. Auch hier ergab die Stichprobenprüfung (jeweils 3 Datensätze aus Item Type Video, Image, Audio) keine Textalternativen. Weiterhin ist keine Anpassungen von Textgröße oder Farbschema möglich und die Suche ist nicht mit Tastatur allein bedienbar. Die Suchfunktion bietet keine Eingabehilfe oder Fehlerkorrektur. Zusätzlich fand die Browsererweiterung aXe 14 Verstöße gegen die Barrierefreiheit in der Anzeige von Suchergebnissen (siehe Screenshot LSE012_Accessibility-Suche) und 4 Verstöße auf Datensatzebene (siehe Screenshot LSE013_Accessibility-Datensatz), in beiden Fällen vor allem wegen fehlendem Farbkontrast.

Audiovisuelle Forschungsdaten haben keine Vorschaufunktion, sondern bieten nur den direkten Download (siehe Screenshot LSE001_AnzeigeDatensatz).

Nutzerx die Support zu Repositoriumsaktionen benötigen können die FAQs¹²⁶ konsultieren oder das zuständige Team per e-Mail kontaktieren¹²⁷.

Die Bibliothek der Hochschule bietet einige Beratungsangebote zum FDM¹²⁸, u.a. in den Bereichen Datenmanagementplan, Metadatenstellung, Versionierung und Ablage in Repositorien [LSE Jd], sowie einem nicht-öffentlichen Moodle Kurs zu FDM. Unterstützung kann per e-Mail erbeten werden [LSE Jd]. Allgemein wird auf SHERPA/RoMEO¹²⁹ [LSE Je], und die UK Data Service Recommended Formats und Library of Congress Recommended Format Specifications [LSE Jb] verwiesen.

K5 Organisation und Infrastruktur

Die Anbindung an die Bibliothek passt zum Ziel des LSERO: "LSE Research Online aims to be a complete database of research produced by LSE Staff." [LSE Je].

Die Hochschule wird nicht explizit als Träger benannt, aber die Bibliothek als fachliche Betreuung festgelegt [LSE Je].

Die zuständige Abteilung "Metadata Services" beschäftigt acht Mitarbeiterx, davon arbeiten 4 an der Betreuung des Repositoriums. Dies enthält einx "Senior Library Assistant" und drei "Library Assistants", die aber gleichzeitig anderweitige Katalogisierungsarbeiten übernehmen [Wil17]. Beschrieben werden die Mitarbeiterx als "information professionals" die Katalogisierungs- und Dienstleistungserfahrung haben [LSE Je].

Die informationstechnische Basis von LSERO ist EPrints 3 [LSE Je], zur Erhebung der Statistik wird das Paket EPrints IRStats2 [AM14] eingesetzt. EPrints wird an der University of Southampton entwickelt [LSE Jg] und ist ein offenes System. Weitere Details sind nicht bekannt. In der Analyse der Webseite wurde deutlich, dass die Suche eher langsam arbeitet.

K6 Gebühren und Finanzierung

Es werden keine Aussagen zum Gebührenmodell oder der sonstigen Finanzierung des LSERO getroffen.

5.6. OceanRep GEOMAR Repository

K1 Veröffentlichung und Nachnutzung

OceanRep akzeptiert u.a. Zeitschriftenartikel, Videos und Forschungsberichte [GEO Jd], mischt also FD und Publikationen. Die Anzahl der Uploads ist unbegrenzt [GEO Ja], wichtiger ist, dass die Einträge der Forschungsrichtung entsprechen und korrekte Metadaten vorliegen [GEO Jd]. Als Format wird ungesperrtes PDF empfohlen, daneben werden auch ASCII, Postscript

¹²⁶ <http://eprints.lse.ac.uk/faq.html>

¹²⁷ <http://eprints.lse.ac.uk/contact.html>

¹²⁸ <http://www.lse.ac.uk/library/research-support/research-data-management>

¹²⁹ <http://www.sherpa.ac.uk/romeo/index.php>

und HTML akzeptiert [GEO Ja]. Eine Übersicht der Publikationstypen¹³⁰ steht Online zur Verfügung und zeigt diverse weitere Arten von Publikationen an. Der Upload ist für Nutzerx selbst oder über die Mitarbeiterx der Bibliothek möglich. Ein Import aus externen Quellen kann über Pubmed ID und XML, DOIs und XML Dateien erfolgen [GEO Jb].

Nach dem Upload werden die Daten manuell validiert, dabei werden u.a. die Metadaten in den Pflichtfeldern, Fehler im Text und Urheberrechtsstatus geprüft [GEO Ja].

Zur Unterstützung bei der Vergabe der Metadaten ist ein Autorxverzeichnis hinterlegt, das den vollen Namen und den Forschungsbereich des Forscherx populiert [GEO Jb]. Der Upload erfolgt über ein Formular mit fünf getrennten Eingabemasken, z.B. für die Eingabe von Dokumententyp oder Schlagworten. Pflichtfelder sind mit Sternen markiert [GEO Jb]. Für die Schlagworte ist eine Systematik der Themengebiete hinterlegt, diese wird von der Bibliothek gepflegt (ebd.). Die Verlinkung auf eine Datenpublikation kann über das Feld "Offizielle URL" erstellt werden (ebd.).

Nach Upload wird auf Datensatzebene ein Zitationsvorschlag angezeigt (siehe Screenshot ORE001_AnzeigeDatensatz), eine explizite Zitationsrichtlinie gibt es nicht. Um die Datensätze zitierfähig zu machen werden URNs vergeben [GEO Ja], außerdem wird das DOI System genutzt (siehe Screenshot ORE001_AnzeigeDatensatz).

Der Screenshot zeigt auch, dass die Landingpages die gewählten Lizenzen anzeigen. Für die offene Lizenzierung wird Creative Commons angeboten [GEO Ja]. Wenn ein Forscherx dies nicht möchte, zeigt das Feld "Lizenz" standardmäßig "Nicht spezifiziert" an [GEO Jb]. Die OceanRep Metadaten können unter Herkunftsangabe für nicht-kommerzielle Zwecke verwendet werden [GEO Jc], sie sind also nicht bindungslos.

Ein Verweis auf zugehörige Publikationen ist über das Feld "Offizielle URL" möglich [GEO Jb], bzw. siehe Screenshot ORE004_Verlinkung. Zugehörige Projekte werden über "Additional Affiliation" zugeordnet oder über "Verwandte URLs" verknüpft (ebd.). Wenn bei Pangaea ebenfalls Daten liegen werden beide Einträge automatisch verknüpft [GEO15], siehe auch Screenshot ORE013_VerlinkungProjekte.

Jede Landingpage zeigt Downloads pro Tag und Monat (siehe Screenshot ORE002_Statistik) und bietet eine erweiterte Anzeige, die z.B. auch monatliche Seitenaufrufe oder die Top Ten Suchbegriffe im Repositorium anzeigt (siehe Screenshot ORE003_MoreStatistics).

Um Forscherx zum Open Access Upload ihrer FD anzuregen werden die erhöhte Zitation und Findbarkeit angeführt [GEO Ja], eine Anzeige der meist heruntergeladenen Paper¹³¹ ist als weiterer Ansporn auf der Webseite. Zur Identifikation von Autorx werden neben ORCID auch GND, Scopus ID und ResearcherID angeboten [Sch16].

Automatische Reports werden im GEOMAR z.B. für die Erstellung von Jahresberichten angeboten, als Beispiele werden Listen für "LOM (Leistungsorientierte Mittelvergabe) sowie POF (Programmorientierte Förderung)" ([Sch14, S. 59]) genannt.

K2 Zugriff und Auffindbarkeit

Zum Upload sind nur Mitarbeiterx des GEOMAR Institut berechtigt [GEO Ja], bzw. beauftragte Personen [GEO Jc].

Der Zugang zu hochgeladenen Daten kann öffentlich (Open), für registrierte Nutzerx (Restricted) oder nur für OceanRep Mitarbeiterx (Closed) eingeschränkt werden. Letzteres ist gedacht für Privatgebrauch, Daten die nicht veröffentlicht werden dürfen oder Daten die vor der Veröffentlichung begutachtet werden sollen [GEO Jb]. Nutzerx können die Metadaten in jedem Fall einsehen (ebd.). Ein zeitlich begrenztes Embargo ist ebenfalls möglich, dazu wird das Feld "Enddatum der Sperre" belegt. Nach Ablauf endet das Embargo automatisch [GEO

¹³⁰<http://oceanrep.geomar.de/view/types/public/>

¹³¹http://oceanrep.geomar.de/cgi/irstats.cgi?page=dashboard&IRS_epchoice=All

Jb]. Über die Erweiterte Suche ist eine gezielte Suche nach Zugangsrechten möglich, darunter auch Embargoed access (siehe Screenshot ORE008_Zugangsrechte).

Für die Recherche stehen drei Einstiege zur Verfügung: Eine Schnellsuche, eine einfache Suche und eine erweiterte Suche. Boolesche Suchoperatoren können benutzt werden [GEO Jb]. Die attributspezifische Suche wird als "Blättern" bezeichnet, hier stehen u.a. die Bereiche Autor, Forschungsbereich oder Themengebiet zur Verfügung [GEO Jb]. Eine Volltextsuche ist möglich (siehe Screenshot ORE009_EinfacheSuche).

Die Findbarkeit der Einträge in Suchmaschinen wird durch die Vergabe von URNs für die Landingpages unterstützt [GEO Ja].

Metadaten können über eine OAI-PMH Schnittstelle abgerufen werden, zusätzlich kann die Repository Software EPrints die Ausgaben diverser Zitationsprogramme wie z.B. Endnote einspeisen um die Literaturliste ebenfalls abzulegen [GEO Ja]. Die Base URL für OAI-PMH wird auf der Startseite angezeigt¹³².

K3 Transparenz und Qualitätssicherung

OceanRep hat eine explizite Richtlinie¹³³ und verweist außerdem auf die Institutsweite "Richtlinie für wissenschaftliche Veröffentlichungen"¹³⁴, die für Forscherx im Intranet verfügbar ist. Zur Validierung wird in den FAQs festgehalten, dass die Dokumente formal und rechtlich geprüft werden [GEO Ja], in der Richtlinie werden dennoch Datengeberx als endgültig verantwortliche Person benannt [GEO Jc].

Hochgeladene Daten können versioniert werden [GEO Jc]. Die Versionierung geschieht über die Funktion "Neue Version", die den Eintrag dupliziert damit an der Kopie Änderungen vorgenommen werden können [GEO Jb]. Die URN zeigt danach auf die Kopie, also auf die neueste Version [GEO Jd], beide Versionen bleiben aber öffentlich in OceanRep [GEO Jb].

Eine Löschung von Datensätzen ist nicht vorgesehen, der Zugriff kann aber durch Bibliothekarx gesperrt werden [GEO Ja], die Nennung eines Grundes ist dabei erwünscht [GEO Jb]. In der Richtlinie heisst es weiterhin, dass Sperrungen vor allem bei urheberrechtlichen Problemen oder aufgrund von wissenschaftlichem Fehlverhalten vorgenommen werden [GEO Jc]. Nach einer Sperrung verbleiben die Landingpages und Links aktiv (ebd.).

OceanRep übernimmt die Verantwortung für die LZA im Abschnitt Preservation Policy der Repository Policies¹³⁵. Dort heisst es u.a., dass Backups immer nach "Best Practice" vorgenommen werden [GEO Jc]. OceanRep lässt sich ein einfaches Nutzungsrecht übertragen um im Bedarfsfall auch "technische Veränderungen" [GEO Ja] für die LZA durchführen zu können. Die Sicherungsdauer wird als "langfristig" [GEO Ja] bezeichnet, bzw. ist die Aussage: "Items will be retained indefinitely." [GEO Jc]. Eine Nachfolgeregelung ist nicht ersichtlich, es wird nur versichert, dass ein geeignetes Archiv die Daten übernehmen würde [GEO Jc].

OceanRep ist nicht zertifiziert.

K4 Nutzererlebnis und Support

Die Nachnutzung wird durch diverse Exportformate (siehe Screenshot ORE005_Exportformate) erleichtert. Außerdem erhalten die Datengeber ein "Automatisches Update von Publikationslisten auf Projekt-, Schiffs-, und persönlichen Webseiten" [Sch14]. Um die Dateneingabe zu erleichtern, kann diese jederzeit unterbrochen und für später gespeichert werden, das Eingabeformular zeigt nach Auswahl des Publikationstyps nur die benötigten Metadatenfelder an, etwaige Fehler in der Eingabe werden farblich markiert und es können teilausgefüllte Vorlagen angelegt werden [GEO Jb]. Die intuitive Bedienung wird durch das Fehlen einer leeren Suche

¹³² <http://oceanrep.geomar.de/>

¹³³ <http://oceanrep.geomar.de/policies.html>

¹³⁴ <https://www.geomar.de/zentrum/einrichtungen/bibliothek/publikationen-at-geomar/>

¹³⁵ <http://oceanrep.geomar.de/policies.html>

beeinträchtigt, da es dadurch nicht möglich ist alle Datensätze anzeigen zu lassen. Die Darstellungsformen die EPrints anzeigt sind etabliert, aber eher veraltet. Um Einsteigerx durch die ersten Suchen zu helfen hat jedes Suchfeld eine Hilfestellung [[GEO Jb](#)].

Bei der Untersuchung der Barrierefreiheit findet die Browsererweiterung aXe 49 Verstöße gegen die WCA Guidelines (siehe Screenshot ORE010_Accessibility-Suche). Die Hauptprobleme sind fehlender Farbkontrast und fehlende Beschriftung von Bildern und Links. Weitere 16 Verstöße werden auf Datensatzebene angezeigt (siehe Screenshot ORE011_Accessibility-Datensatz), wieder fehlt ein ausreichender Farbkontrast und die Beschriftung von Bildern. Bei der manuellen Betrachtung der Webseite wurde festgestellt, dass die Beschreibung von Videos im Abstract erfolgen kann, siehe Beispieldatensatz¹³⁶, der bei Metadatenexport in DC dem dc:description Feld entspricht. Es gibt keine spezifischen Metadatenfelder für die inhaltliche Beschreibung, Transkripte oder Untertitel für nicht-textuelle Daten. Weiterhin wurden keine Textalternativen gefunden. Die Webseite des GEOMAR Zentrums erlaubt eine Einstellung der Schriftgröße (siehe Screenshot ORE007_Accessibility-Schrift), auf den Unterseiten des OceanRep ist dies aber nicht implementiert. Die Suche ist mit Tastatur allein bedienbar, aber nur umständlich. Es wird keine Eingabehilfe oder Fehlerkorrektur in den Suchfeldern angeboten.

Für Bilder und Videos gibt es eine Vorschaufunktion direkt im Browser, diese ist aus der Liste der Suchergebnisse [[GEO Jb](#)] oder vom Datensatz erreichbar (siehe Screenshot ORE012_Vorschau).

Für den Support zu Repositoriumsaktionen werden eine Suchhilfe und Hilfeseiten für Forscherx¹³⁷, sowie eine Zusammenstellung von FAQs¹³⁸ angeboten. Individuelle Beratung wird durch Bibliotheksmitarbeiterx angeboten [[GEO Jb](#)].

Hinsichtlich aktiver Unterstützung zum FDM können Forscherx ihre Daten wie oben erwähnt durch Bibliothekarx einpflegen lassen, Schmidt sagt dazu: "Der Service [...] wird sehr hoch geschätzt." [[Sch14](#), S. 60]. Weiterhin wird zum Thema Forschungsdaten auf www.forschungsdaten.org verwiesen [[GEO15](#)].

K5 Organisation und Infrastruktur

Die Anbindung an die Bibliothek des GEOMAR passt zu den Zielen von OceanRep, zusätzlich möchte die Bibliothek als Ansprechpartner für wissenschaftliche Publikationen assoziiert werden [[Sch14](#), S. 60].

Die Zuständigkeiten werden auf der Startseite des Repositoriums bekanntgegeben: Die inhaltliche Betreuung erfolgt durch die Bibliothek, die technische durch das Kieler Datenmanagementteam [[GEOoJ](#)]. GEOMAR selbst ist eine Stiftung öffentlichen Rechts [[Sch14](#), S. 58], die Rechtsform von OceanRep ist unbekannt.

Die Webseite liefert keine Hinweise wer an OceanRep mitarbeitet und in welcher Kapazität, aber die Bibliothek hat drei Mitarbeiterx¹³⁹ und das Kieler Datenmanagementteam hat fünf Mitarbeiterx¹⁴⁰.

Als Repository Software wurde EPrints gewählt: "Der große Vorteil von EPrints ist, dass die Software von Wissenschaftlern für Wissenschaftler erstellt wurde und dass sie nicht von einer kommerziellen Firma vertrieben wird." [[Sch14](#), S. 57]).

K6 Gebühren und Finanzierung

OceanRep erhebt keine Gebühren für den Download von Daten [[GEO Jc](#)].

Die ursprüngliche Finanzierung wird auf der OceanRep Startseite erwähnt "OceanRep [...] wurde mitfinanziert vom Exzellenz-Cluster 'Future Ocean!'" [[GEOoJ](#)]. Weitere Details liefert

¹³⁶http://dx.doi.org/10.3289/MATING_CLIONE_LIMACINA_2010

¹³⁷<http://oceanrep.geomar.de/help>

¹³⁸<http://oceanrep.geomar.de/faqs.html>

¹³⁹<https://www.geomar.de/zentrum/einrichtungen/bibliothek/team/>

¹⁴⁰<https://www.geomar.de/de/zentrum/einrichtungen/rz/daten/>

Schmidt in einem Artikel: "Der Cluster sowie beide Sonderforschungsbereiche [574 und 754] entschieden sich zur Beteiligung an dem Repository." [Sch14, S. 60]. Die Schätzung aus 2014 erwartete laufende Kosten von 3000 Euro pro Jahr (ebd.).

5.7. Marine-Geo Digital Library

K1 Veröffentlichung und Nachnutzung

Die MGDL enthält FD aus den Geowissenschaften, darunter solche aus Tiefseemessung und Sonar [Mar Jb]. Diese werden in diversen Dateitypen akzeptiert, eine Liste der Formate¹⁴¹ ist Teil des Vokabulars [Mar Jb]. Für Videos wird das Format .mov empfohlen.

Forscherx können ein auf Google Forms gehostetes Data Registration Form nutzen um Daten zu beschreiben und per Filesharing (z.B. DropBox) einzureichen, oder sich kostenlos bei GeoPass registrieren und das File Uploader Tool verwenden [Mar Jd]. Nach Login bei GeoPass werden Forscherx informiert, dass Dateien maximal 50MB groß sein dürfen und der File Uploader nur ausgewählte Datentypen akzeptiert (ebd., aber nur nach Login sichtbar).

Daten werden grundlegend mit Metadaten beschrieben und optional mit temporalen und geografischen Daten angereichert, weitere disziplinspezifische Felder stehen z.B. für seismische Daten bereit. Außerdem werden für die Beschreibung von Expeditionen Felder wie "Technical Reports", "Funding Information" und "Participants and their Affiliations" angeboten [Mar Jf]. Eine Abfrage über Schnittstellen gibt ISO19115¹⁴² Metadaten aus [Mar Jh]. Um Forscherx bei der Eingabe von Daten zu unterstützen haben einige Felder eine Auto-Vervollständigung [Mar Jc].

Beide Upload-Möglichkeiten sind formularbasiert und offen zugänglich. Das Data Registration Form ist auf Google Forms¹⁴³ einsehbar, der Marine-Geo Digital Library File Uploader erst nach kostenfreier Registrierung (siehe Screenshot MGD003_UploadForm). Beide weisen auf Pflichtangaben hin: Im Google Form sind dies z.B. Data Types und Data Processing Level, im Uploader sind es z.B. Investigator und Expedition or Compilation ID. Die Submission Guideline erwähnt u.a. auch "Funding agency and award numbers" [Mar Jb].

Hinsichtlich der Dokumentation der FD wird darauf verwiesen, dass alle NSF-geförderten Projekte "open and rapid access to quality-controlled and fully documented data" [Nat16] leisten müssen, dies betrifft einen Großteil der MGDL FD.

Die Verlinkung auf eine anderweitige Datenpublikation ist über das "Related Reference(s)" Metadatum im MGDL File Uploader möglich (siehe Screenshot MGD003_UploadForm), außerdem werden Rohdaten und darauf basierende Datenprodukte verknüpft [Mar Jf].

Die Zitationsrichtlinie des MGDL ist in den Terms of Use enthalten [Mar Jj], zusätzlich wird auf Datensatzebene im Reiter "Data Citation Information" ein Zitationsvorschlag gegeben (siehe Screenshot MGD006_Citation). Zitierfähig werden die Datensätze über die DOI (siehe Screenshot MGD005_Datensatzansicht), die Vergabe erfolgt direkt über das FDR [Mar Je].

Bedingung für die Nachnutzung der Daten ist, dass Datengeber und das Marine Geoscience Data System zitiert werden [Mar Jj]. Die Daten und Metadaten sind CC-BY-NC-SA 3.0 lizenziert [Mar Ja] und schließen damit kommerzielle Nutzung aus.

Forschungsdaten und deren zugehörige Publikationen werden im Datenmodell verknüpft [Mar

¹⁴¹http://www.marine-geo.org/tools/search/vocab.php?use_is_displayed=T&vocab=vocab_file_format_id

¹⁴²Deutsche Übersetzung von ISO19115 - Geographic Information auf Geoportal.de: http://www.geoportal.de/SharedDocs/Downloads/DE/GDI-DE/Deutsche_Uebersetzung_der_ISO-Felder.html

¹⁴³<https://docs.google.com/forms/d/e/1FAIpQLSf5Uu0eBZ71rjb29WivA3vFuFHYy29DobxLWXFSM0fCbJX0iw/viewform>

[Jf], die Anzeige auf der Webseite ist in Screenshot MGD007_References abgebildet. Außerdem bestehen Kooperationen mit Zeitschriften aus den Geowissenschaften um die Verlinkung zu unterstützen [Mar Je]. Das Datenmodell sieht außerdem vor, dass Daten aus Projekten untereinander verknüpft werden [Mar Jf], siehe Screenshot MGD008_RelatedDataSets. Das FDR zeichnet diverse Statistiken auf Seitenebene auf. Eine Übersicht befindet sich auf der Webseite¹⁴⁴, darunter auch eine Auflistung aller Zitationen der FD¹⁴⁵. Datengeber erhalten "semi-annual data download reports [...] to inform them of downloads/views of their data." [Mar Je].

K2 Zugriff und Auffindbarkeit

Daten können öffentlich, unter "proprietary hold" oder passwortgeschützt in die MGDL gestellt werden. Dabei bezieht sich "proprietary hold" auf Daten aus Projekten mit eigenen Nutzungsbedingungen für die FD [Mar Jj] und Passwortschutz ist für Daten vor der Veröffentlichung bestimmt [Mar Jd]. Ein Embargo ist ebenfalls über den "proprietary hold" möglich, die Sperrfrist wird auf Datensatzebene angezeigt (siehe Screenshot MGD009_Embargo).

Zum Upload sind alle Nutzerx berechtigt, angenommen werden aber nur "sensor-based geoscience data" [Mar Jd]. Der Großteil der Daten im FDR stammt von Forschungsschiffen und sonstigen Forschungseinrichtungen. Nutzerx die diese Daten nachnutzen können darauf basierende Ergebnisse wieder in der MGDL einreichen: "[...] in particular processed and interpreted data products supporting publications and field data that has not been archived." [Mar Je].

Die Recherche im MGDL erfolgt direkt auf der Webseite oder über den IEDA Integrated Data Browser¹⁴⁶. Das lokale Angebot erlaubt eine Suche über den gesamten Datenbestand oder über einzelne sog. "Community Data Portals", die Themen- oder Projektspezifisch zusammengestellt sind [Mar Jg]. Eine Freitextsuche ist über den IEDA Browser möglich (siehe Screenshot MGD011_IEDAIntegratedDataBrowser), direkt im FDR kann nur attributspezifisch über Auswahllisten gesucht werden, z.B. über data type oder scientist name [Mar Ji]. Die Suche kann über eine Karte (siehe Screenshot MGD010_Kartensuche) oder Textanzeige begonnen werden, beide bieten diverse Suchfelder die alle mit einem Vokabular verknüpft sind (siehe Screenshot MGD012_Vokabulare). Der IEDA Browser erlaubt eine facettierte Suche (siehe Screenshot MGD011_IEDAIntegratedDataBrowser).

Zu indexierbaren Landingpages wird keine Aussage getroffen und eine Known-Item Search für einen Beispieldatensatz war nicht erfolgreich.

Die REST API "MGDS FileServer" erlaubt die Abfrage von Metadaten auf Datensatz- und Datenebene [Mar Jh], eine OAI-PMH Schnittstelle wird nicht erwähnt. Die APIs "MGDS FileDownloadServer" und "MGDS ArchiveDownloadServer" erlauben den Download von einem oder mehreren Datensätzen. Zusätzlich gibt es eine OGC Schnittstelle entsprechend den Standards des Open Geospatial Consortium und eine MGDS Search API für die Suche (ebd.).

K3 Transparenz und Qualitätssicherung

Das FDR hat Terms of Use¹⁴⁷ und Data Submission Guidelines¹⁴⁸.

Zu Datenvalidierung, Versionierung und Löschung werden keine Aussagen getroffen. Zur Langzeitarchivierung gibt es nur eine allgemeine Aussage: "The Marine-Geo Digital Library offers data preservation services including long-term archiving [...]." [Mar Jg].

Die MGDL ist nicht zertifiziert.

K4 Nutzererlebnis und Support

Eine Maßnahme um die Usability zu erhöhen sind die verschiedenen Sucheinstiege: Neben der

¹⁴⁴http://www.marine-geo.org/about/statistics_holdings.php

¹⁴⁵<http://www.marine-geo.org/about/citations.php>

¹⁴⁶<http://catalog-dev.iedadata.org:8080/geoportal/>

¹⁴⁷http://www.marine-geo.org/about/terms_of_use.php

¹⁴⁸<http://www.marine-geo.org/submit/guidelines.php>

Kartensuche gibt es auch ein "text-only interface (for slow connections)" (siehe Screenshot MGD010_Kartensuche). Unter dem Reiter News¹⁴⁹ werden analog einer Neuerwerbungsliste die neueste Einträge in der MGDL angezeigt.

Die Bedienung der Kartensuche ist eher nicht intuitiv, generell werden aber etablierte Darstellungsformen verwandt. Die eingeschränkte Suchhilfe, vor allem zum Thema Suchoperatoren, erschwert die Suche nach Daten.

Die Webseite ist nicht barrierefrei, die Browsererweiterung aXe findet 32 Verstöße gegen Barrierefreiheit in den Suchergebnissen (siehe Screenshot MGD013_Accessibility-Suche) und 15 Verstöße auf Datensatzebene (siehe Screenshot MGD014_Accessibility-Datensatz). Als Eingabehilfe ist eine Auto-Vervollständigung für jedes Suchfeld vorhanden [Mar Jh].

FD können in diversen Anwendungen visualisiert werden, darunter Virtual Ocean, die GeoMapApp, und der NASA World Wind 3-D earth browser, sowie weitere je nach Art der Daten, z.B. für seismische Messungen [Mar Jg]. Eine Vorschaufunktion gibt es nur für Bilder (siehe Screenshot MGD015_Vorschau).

Zur Unterstützung der Nutzerx wurden die Webseiten "Frequently-Asked Questions about data terminology"¹⁵⁰, Suchhilfen¹⁵¹, Übersicht über Programmierschnittstellen¹⁵² und File Uploader - Tips and Help¹⁵³ zusammengestellt. Außerdem gibt es Videotutorials und ein Webinar zu unterschiedlichen Funktionalitäten¹⁵⁴. Supportanfragen können über ein Kontaktformular¹⁵⁵ an das MGDS Team gestellt werden. Die Terms of Use fordern Nutzerx außerdem explizit dazu auf Fehler in den Datensätzen zu melden [Mar Jj].

Eine Beratung zum Forschungsdatenmanagement findet nicht statt.

K5 Organisation und Infrastruktur

Die MGDL wird gehostet im Lamont-Doherty Earth Observatory der Columbia University [Lam J], die fachliche Anbindung an das Institut ist für das FDR angebracht. Der Träger wird sichtbar benannt: "Funding for MGDS is provided by the National Science Foundation as part of IEDA Cooperative Agreement." [Mar Ja], zusätzlich erfolgt ein Hinweis im Footer der auf jeder Webseite zu sehen ist.

Das Team im MGDS das für das FDR zuständig ist, umfasst 9 Mitarbeiterx und 2 Projektleiterx [Lam J]. Die Qualifikationen weisen die Personen als Fachwissenschaftler bzw. Informatiker aus, Informationswissenschaftler sind nicht genannt.

Zu den Forschungs- und Entwicklungstätigkeiten der zuständigen Abteilung liegen keine Hinweise vor.

Zur technischen Basis wird nur erwähnt, dass im Backend eine relationale Datenbank läuft, die auf die Bedürfnisse der betreuten Communities ausgerichtet ist [Mar Jf]. Zu den Ladezeiten wird gewarnt, dass die geografische Suche mehr Zeit in Anspruch nehmen kann als die einfache Textsuche [Mar Ji].

K6 Gebühren und Finanzierung

Es werden keine Aussagen zu Gebühren oder der Finanzierung des FDR getroffen.

¹⁴⁹ <http://www.marine-geo.org/portals/antarctic/news.php>

¹⁵⁰ http://www.marine-geo.org/help/data_FAQ.php

¹⁵¹ http://www.marine-geo.org/help/search_results_help.php und http://www.marine-geo.org/help/search_results_help.php

¹⁵² http://www.marine-geo.org/tools/web_services.php

¹⁵³ http://www.marine-geo.org/help/contribute_help.php

¹⁵⁴ <http://www.marine-geo.org/help/index.php>

¹⁵⁵ <http://www.marine-geo.org/about/contact.php>

5.8. DR-NTU Data Repository

K1 Veröffentlichung und Nachnutzung

Das FDR der Nanyang Universität akzeptiert alle Dateitypen und legt keine Maximalgröße für Dateien fest, die technischen Grenzen der Verbindung liegen aber bei 3,5 GB pro Upload [NTU17h]. Der Upload erfolgt über ein Formular [NTU17k] oder "Tabular Ingest" für tabellarische Daten aus Excel oder SPSS [NTU17m]. Eine Datenintegritätsprüfung findet nicht statt [NTU17i].

Die Metadaten werden organisatorisch getrennt in Citation, Domain specific und File-level Metadaten [NTU17l]. Eine volle Liste der Schemata die verwendet werden ist unter Supported Metadata and References¹⁵⁶ Online zugänglich, für Citation Metadaten sind dies u.a. DDI Lite, und Dublin Core und domänenspezifische wie das VOResource Schema. Insgesamt werden etwa 100 Metadatenfelder unterstützt [NTU17c]. Die Eingabe erfolgt über ein Formular und Pflichtangaben sind markiert [NTU17k]. Einzelne Felder sind mit Vokabularen hinterlegt, für Sprachen z.B. die ISO 639-2 Codes, sowie fachspezifische Ontologien [NTU17l].

Den FD zugehörige Dokumentation und Software wird explizit mit in die zu archivierenden Objekte einbezogen [NTU17a].

Die Richtlinien zur Zitation von FD richtet sich nach den Dataverse Community Norms und verpflichten Nachnutzerx die Datengeberx zu zitieren: "These citations include the data authors, data identifier, and other information in accordance with the Joint Declaration of Data Citation Principles [...]" [NTU17e]. Zur Sicherstellung der Zitierfähigkeit sind alle Datensätze mit DOI versehen [NTU17h].

Die Nutzungsbedingungen werden in den Terms of Use festgelegt und verweisen auf eine CC-BY-NC Lizenz, bzw auf die Einhaltung eines "Custom Data Usage Agreement" ([NTU17i, S. 7]), das vom Datengeberx festgelegt wird. CC-BY-NC ist die voreingestellte Standardlizenz [NTU17h].

Die Verlinkung auf eine zugehörige Publikation erfolgt in den Metadaten [NTU17d], siehe dazu auch Screenshot NTU002_Verlinkung. Die Architektur der Repository Software Dataverse sieht für Projekte die Zuordnung über die gleichnamigen Dataverses vor.

Ebenfalls in der Software enthalten ist die Erhebung von Metriken über Google Analytic und Piwik [NTU17g]. Für die Umsetzung in der Praxis siehe Screenshot NTU004_Metriken.

Forscherx werden über den Verweise auf erhöhte Zitationsraten motiviert ihre FD zu veröffentlichen [NTU Ja]. Die Integration von Autoreneidentifikation per ORCID ist noch in Planung [NTU17g].

Das FDR sammelt die Downloadzahlen und stellt diese dem Datengeberx zur Verfügung [NTU17i].

K2 Zugriff und Auffindbarkeit

Datensätze sind standardmäßig öffentlich verfügbar, bei Zugriffsbeschränkung werden Nutzerx gebeten die "Terms of Access" festzulegen ([NTU17f, S. 3]).

Zum Upload sind alle Universitätsangehörigen berechtigt inkl. den Studierenden [NTU17h].

Für den Sucheinstieg stehen einfache, erweiterte Suche und Browsing zur Verfügung. Die erweiterte Suche erlaubt die Suche auf Sammlungsebene (Dataverse) und in den Metadaten auf Datensatzebene und Forschungsdatenebene [NTU17j]. Die möglichen Suchoperatoren sind in der Suchhilfe nicht enthalten. Ein Test mit Trunkierung und exakter Suche war erfolgreich, Wildcards werden nicht unterstützt (siehe Screenshot NTU006_Wildcard). Die attributspezifische Suche ist in der erweiterten Suche enthalten und Facettierung ist möglich (ebd.).

¹⁵⁶https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/DRNTUData_SupportedMetadataReferences-2e09yn2.pdf

Die Landingpages werden von Suchmaschinen wie Google und Bing indexiert, dies kann durch Nutzerx für ein Dataverse oder einen Datensatz erlaubt oder verhindert werden ([NTU17i, S. 1]).

Für die Abfrage der Metadaten steht eine OAI-PMH Schnittstelle in der Dataverse Software zur Verfügung [NTU17g], aber eine Base URL zur Abfrage wird auf der Webseite nicht angezeigt. Ebenso technisch möglich sind die Implementierung einer SWORD API, Search API, Data Access API und selbstentwickelter APIs [NTU17g], aber zur tatsächlichen Benutzung im DR-NTU Repository werden keine Aussagen getroffen.

K3 Transparenz und Qualitätssicherung

Das FDR hat eine DR-NTU (Data) Policy¹⁵⁷ und die DR-NTU (Data) General Terms of Use¹⁵⁸. Die Universität hat zudem eine allgemeine NTU Research Data Policy¹⁵⁹. Weiterhin ist ein Sample Data Usage Agreement¹⁶⁰ öffentlich zugänglich.

Alle Änderungen an veröffentlichten Datensätzen führen automatisch zu einer Versionierung, auch Änderungen an den Metadaten ([NTU17f, S. 9]). Welche Änderung wann vorgenommen wurde wird aufgezeichnet und ist auf Datensatzebene sichtbar (ebd., S. 10), siehe dazu auch Screenshot NTU003_Versionierung.

Um Datensätze aus dem FDR zu entfernen wird eine Deaccession vorgenommen, dies entspricht einer Sperrung und wird bezeichnet als "[...] a very serious action that should only occur if there is a legal or valid reason for the dataset to no longer be accessible to the public." [NTU17f, S. 10]. Die Landingpage ist weiterhin unter der DOI erreichbar, die Daten und Metadaten werden jedoch entfernt (ebd.).

Für die LZA wird in den FAQ die Verantwortung übernommen: "All data deposited in DR-NTU (Data) will be preserved and archived [...]" [NTU17h]. Als Maßnahmen werden die Anfertigung von Kopien und Übertragung in andere Formate erwähnt [NTU17i]. Die NTU Research Data Policy verpflichtet Universitätsangehörige zur Aufbewahrung von FD für 10 Jahre [Nan15]. Ob das FDR dies garantiert wird nicht deutlich.

DR-NTU (Data) ist nicht zertifiziert, das FDR wurde aber auch wie im vorigen Kapitel erwähnt erst im September 2017 eröffnet [Wen17].

K4 Nutzererlebnis und Support

Die Bedienung des FDR ist intuitiv, wird aber durch eine eingeschränkte Suchhilfe beeinträchtigt. Die Webseite ist zudem nicht barrierefrei, z.B. findet die Browsererweiterung aXe 65 Verstöße gegen die WCAG in den Suchergebnissen (siehe Screenshot NTU007_Accessibility-Suche) und 22 Verstöße auf Datensatzebene (siehe Screenshot NTU008_Accessibility-Datensatz), in beiden Fällen vor allem wegen niedrigem Farbkontrast. Eine inhaltliche Beschreibung von Videos kann zwar per Freitext in Description umgesetzt werden (Praxisbeispiel: Forscherx legt Textdatei mit Transkript bei, siehe Screenshot NTU009_BeispielTranskript), dies ist aber nicht normiert.

Für die Datenvisualisierung bietet die Softwareplattform die Funktion WorldMap für Geodaten [NTU17b, S. 3], sowie TwoRavens für tabellarische Daten (ebd. S. 1), ob diese Funktion für das DR-NTU Repository umgesetzt wurde ist unklar.

Das FDR bietet eine Reihe von Hilfsdokumenten¹⁶¹. Die Übersichtsseite gibt an, dass die User Guides von den Dataverse Guides adaptiert wären. Es scheint aber in weiten Teilen direkter

¹⁵⁷<http://blogs.ntu.edu.sg/lib-datamanagement/dr-ntu-data-policies/>

¹⁵⁸https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/12/DRNTUData_GeneralTermsOfUse_v2_01122017-2azwwrc.pdf

¹⁵⁹<http://research.ntu.edu.sg/riero/RI/Pages/Research-Data-Policies.aspx>

¹⁶⁰<https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/Sample-Data-Usage-Agreement%2D17m4cli.pdf>

¹⁶¹<http://blogs.ntu.edu.sg/lib-datamanagement/dr-ntu-data-user-guides/>

Wortlaut des Dataverse Projektes¹⁶² zu sein. Verfügbar sind ein Upload Guide¹⁶³, User Guides zu diversen Themen¹⁶⁴ wie z.B. Finding and Using Data oder Supported Metadata and References, sowie FAQ¹⁶⁵. Schulungsangebote wurden mit Stand September 2017 angekündigt [Wen17].

Supportanfragen werden über ein Kontaktformular¹⁶⁶ entgegengenommen.

Zur FDM- Beratung wird die Bibliothek in der Research Data Policy verpflichtet [Nan15]. Angeboten werden u.a. eine Einführung in das FDM und DMP Workshops¹⁶⁷, sowie Hilfe bei der Metadatenerstellung, Dokumentation von FD und Beratung zu Open Access [Aso17]. Praktische Hilfe wird in Form eines "mediated deposit service" [Wen17] und Vorlagen für DMP¹⁶⁸ angeboten. Die Bibliothek verweist außerdem auf diverse Standards, z.B. die RCUK Common Principles on Data Policy [NTU Ja], die Harvard Dataverse Policies¹⁶⁹, re3data und eine Anzahl von Datenjournalen¹⁷⁰.

K5 Organisation und Infrastruktur

Das FDR wird von der Bibliothek fachlich und von der IT-Abteilung der Universität technisch betreut. Die enge Anbindung ist dem gesetzten Ziel das "institutional open access research data repository for Nanyang Technological University" [NTU17a] zu sein angemessen.

Zur Personalausstattung liegen kaum Informationen vor und zur Qualifizierung keine. Auf der About Us Seite des Research Data Management Team werden 3 Mitarbeiterx angezeigt [NTU Jb], zu den Mitarbeiterx des CITS wird keine Aussage getroffen. Ebenso fehlen Hinweise zu den Forschungs- und Entwicklungstätigkeiten.

An technischen Details ist wenig bekannt, auf den Webseiten des FDR heisst es nur, dass DR-NTU (Data) "protected and secure" [NTU17h] ist und dass das FDR auf dem von der University of Harvard entwickelten Dataverse basiert [NTU17a].

K6 Gebühren und Finanzierung

Es fallen keine Gebühren für den Upload von FD in DR-NTU (Data) an [NTU17h]. Zur Finanzierung werden keine Aussagen getroffen.

¹⁶²<http://guides.dataverse.org/en/4.3/user/>

¹⁶³https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/DRNTUData_HowToUpload-2bmgd1.pdf

¹⁶⁴<http://blogs.ntu.edu.sg/lib-datamanagement/dr-ntu-data-user-guides/>

¹⁶⁵<http://blogs.ntu.edu.sg/lib-datamanagement/dr-ntu-data-faq/>

¹⁶⁶<https://blogs.ntu.edu.sg/lib-datamanagement/contact-us/>

¹⁶⁷<https://blogs.ntu.edu.sg/lib-datamanagement>

¹⁶⁸<http://blogs.ntu.edu.sg/lib-datamanagement/dmp/ntu-data-management-plan>

¹⁶⁹<https://dataverse.org/best-practices/harvard-dataverse-policies>

¹⁷⁰<https://blogs.ntu.edu.sg/lib-datamanagement/wheretoshare/>

6 Vergleich der Forschungsdatenrepositorien

Auf Basis der Ergebnisse der Analyse die im vorigen Kapitel beschrieben wurden erfolgt in diesem Kapitel ein Vergleich aller FDR. Die Grundlage bilden die ausgefüllten Kriterienkataloge in Anhang A.2 und ein abstrahierter Vergleich in Tabellenform in Anhang A.3. Unterteilt ist das Kapitel in die sechs Kategorien der Erfolgskriterien nach Steinhof [Ste17, S.13ff].

6.1. K1 Veröffentlichung und Nachnutzung

K1.1 Datenabgabe

Fünf der betrachteten FDR akzeptieren alle Dateiformate. OceanRep nimmt laut FAQ nur PDF, Postscript, HTML, und ASCII an [GEO Ja], die Übersicht der Publikationsarten¹⁷¹ widerspricht dieser Aussage. LSERO gibt keine Auskunft, die Bibliothek pflegt jedoch eine Liste mit empfohlenen offenen Formaten¹⁷². Ähnliche Listen werden von Phaidra¹⁷³, MGDL¹⁷⁴ und SDR¹⁷⁵ zusammengestellt. Zenodo akzeptiert auch Formate die "preservation unfriendly" [Zen Jd] sind.

Sechs FDR ermöglichen den Upload von Forschungsdaten über den Browser, bis auf DR-NTU gibt es jeweils mindestens eine andere Option. In drei Repositorien können Daten von den Mitarbeiterx der Repositorien direkt eingespeist werden. Phaidra bietet die Software Phaidra-Importer¹⁷⁶ und OceanRep den Import über Identifier wie DOI oder Pubmed ID [GEO Jb]. Zenodo bietet den Upload über REST Schnittstelle an [Zen Jh]. Wie Daten in BDR eingespeist werden ist unbekannt, ebenso wie Forscherx selbst Daten in LSERO hochladen.

Datenintegrität wird nur im BDR, LSERO und OceanRep geprüft. In allen drei Fällen geschieht dies durch Bibliothekarx, automatisierte Prüfungen wurden nicht genannt. DR-NTU und Zenodo führen keine Integritätsprüfung durch. Drei FDR geben keine Hinweise.

K1.2 Metadaten

Vier FDR nennen die Kategorien ihres Metadatenschemas. Phaidra Metadateneditor ist in 9 Module geteilt, darunter Allgemeine und Technische Angaben, Lizenzen und administrative Zuordnungen [Abt08, S. 24]. Das BDR folgt einer ähnlichen Aufteilung und unterscheidet u.a. beschreibende von strukturellen oder rechtebeschreibenden Metadaten [Bro Ji]. Die MGDL teilt in bibliografische und örtliche sowie zeitliche Angaben auf und hat zusätzliche Felder für Angaben zu Forschungsexpeditionen [Mar Jf]. DR-NTU trennt vor allem domänenspezifische Metadaten von allgemeinen Angaben [NTU17i].

LSERO und OceanRep geben nicht öffentlich an ob sie ein standardisiertes Metadatenschema verwenden. MGDL hat ein ISO 19115-basiertes Schema und Zenodo bezieht sich auf das JSON-Schema. Vier weitere geben mehrere Vorbilder an, in allen ist Dublin Core vertreten. Das SDR nutzt zusätzlich MODS, Phaidra LOM und METS, das BDR MODS und METS und

¹⁷¹<http://oceanrep.geomar.de/view/types/public/>

¹⁷²<http://www.lse.ac.uk/library/research-support/research-data-management/file-formats>

¹⁷³<https://backend.univie.ac.at/index.php?id=23004&L=0>

¹⁷⁴http://www.marine-geo.org/tools/search/vocab.php?use_is_displayed=T&vocab=vocab_file_format_id

¹⁷⁵<https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-formats>

¹⁷⁶<https://phaidra.univie.ac.at/view/o:471960>

DR-NTU hat eine Liste von nachgenutzten Metadatenschemata¹⁷⁷, die allgemeine und domänenspezifische beinhaltet. Insgesamt hat keines der FDR das vollständige Metadatenschema offengelegt.

Zur Unterstützung bei der Metadatenvergabe bieten OceanRep und MGDL eine Autovervollständigung ausgewählter Felder an, LSERO plant dies für die Zukunft. Phaidra und OceanRep erlauben das Speichern einer Metadatenvorlage und Phaidra bietet außerdem eine Ingest-Applikation [Gan17, S. 109]. Das SDR empfiehlt eine Reihe externer Werkzeuge¹⁷⁸. Die übrigen drei FDR nennen keine Hilfs-Werkzeuge.

Sechs FDR bieten ihren Nutzerx einen formularbasierten Upload an, in denen auf Pflichtangaben hingewiesen wird. Die Form der Eingabe von Metadaten für BDR und LSERO ist unbekannt.

Als Unterstützung für die Schlagwortvergabe werden von LSERO und OceanRep nicht weiter spezifizierte kontrollierte Vokabulare angeboten. OceanRep pflegt das Vokabular selbst [GEO Jb]. Phaidra bietet verschiedene Klassifikationen und Thesauri wie ÖFOS und Getty [Abt Jf], DR-NTU hat einzelne Felder mit ISO-Codes (z.B. für Länder) oder fachspezifischen Ontologien und Taxonomien ausgestattet [NTU17]. Das SDR lässt nur freie Schlagworte vergeben ([Fro13, Minute 08:00]) und BDR und MGDL treffen keine Aussage.

K1.3 Dokumentation

Etwaige Dokumentation, die hochgeladene FD tiefer beschreibt, ist im SDR, BDR, bei Zenodo und im DR-NTU als Datei im Datensatz inbegriffen. Das SDR erlaubt außerdem die Verlinkung einer extern abgelegten Dokumentation als "related item" ([Fro13, Minute 9:20]). Phaidras Projekt "code4research" thematisiert die verwandte Archivierung von Software aus Forschungsprojekten [Abt Jd], äußert sich aber nicht zum aktuellen Umgang mit FD-Dokumentation. In der MGDL ist der Upload einer Dokumentation grundsätzlich über "Related Reference(s)" möglich, wird aber nicht explizit für diesen Zweck bereitgestellt. Eine Verlinkung auf Datenpublikationen ist bei Zenodo und in der MGDL explizit vorgesehen und wird im SDR, BDR und in OceanRep über vorhandene Metadatenfelder ermöglicht, die aber für andere Verlinkungen vorgesehen sind. Die übrigen drei FDR äußern sich nicht zu Datenpublikationen.

K1.4 Zitierfähigkeit

Sieben FDR zeigen auf Datensebene einen Zitationsvorschlag an, davon folgen Zenodo und DR-NTU den Data Citation Principles und MGDL hat eine eigene Zitationsrichtlinie in den Terms of Use [Mar Jj]. Nur Phaidra gibt keine Hinweise wie FD für die Nachnutzung zitiert werden sollen. Ebenfalls sieben der untersuchten Repositorien verwenden DOI als persistente Identifier, lediglich das SDR nutzt PURLs mit DRUID für die Objektidentifizierung [Fro15]. Zusätzlich zu den DOI nutzt Phaidra als Identifier für die Landingpages das Handle-System [Abt Jf] und OceanRep nutzt URN [GEO Ja].

K1.5 Lizenzmodelle

Die Nutzungsbedingungen für die FD werden in den Richtlinien (Zenodo, BDR), in den Terms of Use (MGDL, DR-NTU) oder im Fall von Phaidra in beidem festgelegt. OceanRep tut dies im Rahmen der Suchhilfe und LSERO in den FAQ. Das SDR legt keine globalen Nutzungsbedingungen fest. Alle untersuchten Repositorien bieten offene Lizenzierung an, dies beinhaltet überall mindestens die Creative Commons Lizenzen. Zusätzlich bietet das SDR noch ODC [Sta Jb], das BDR die GNU GPL (siehe Screenshots BDR011_Lizenz-GNU) und DR-NTU bietet ein sogenanntes Custom Data Usage Agreement [NTU17i]. Datenschutzrichtlinien werden ent-

¹⁷⁷https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/DRNTUData_SupportedMetadataReferences-2e09yn2.pdf

¹⁷⁸<https://library.stanford.edu/research/data-management-services/data-best-practices/creating-metadata/metadata-tools>

weder vom Träger übernommen (LSE, DR-NTU) oder in den Nutzungsbedingungen erwähnt (Phaidra, Zenodo). Die übrigen vier FDR machen keine Angaben zum Datenschutz. Ob die Metadaten lizenziert sind geben nur vier Repositorien an, davon zwei bindungslos (Phaidra, Zenodo) und zwei mit anderen Lizenzen: OceanRep MD dürfen nur nicht-kommerziell und mit Quellenangabe genutzt werden [GEO Jc] und MGDL lizenziert die MD als CC-BY-NC-SA [Mar Ja].

K1.6 Verweise auf Veröffentlichungen

Sieben untersuchte Repositorien sehen die Verknüpfung von FD und zugehörigen Publikationen vor. Lediglich das BDR trifft dazu keine explizite Aussage, über das Metadatum "Related Item" ist eine Verlinkung dennoch möglich. Die Verlinkung von FD und Publikationen, die diese nachnutzen ist im Gegensatz dazu auf keiner Webseite erwähnt. Mindestens ein respektives Metadatenfeld ermöglicht die Verlinkungen von nachnutzenden Publikationen aber für das SDR ("Related Content"), Zenodo ("Related/alternate identifiers") und das BDR ("Related Item"). Alle analysierten Repositorien sehen die Verknüpfung mit zugehörigen Projekten oder anderen Forschungsdaten vor. Dies geschieht bei Zenodo, BDR und LSERO über Sammlungen, die "Related Content" Funktion im SDR, den Forschungsbereich in OceanRep, die Dataverses bei DR-NTU oder sonstige unspezifizierte Funktionen (Phaidra, MGDL). Zenodo bietet zusätzlich Communities, die von Nutzerx je nach Bedarf erstellt werden können.

K1.7 Nutzungsstatistik

Fünf Repositorien erheben Nutzungsstatistiken. Die von MGDL werden jedoch nur Datengeberx zur Verfügung gestellt [Mar Je]. Das SDR erhebt bisher keine Statistiken, formuliert es aber als "not yet" [Sta Je] und das BDR erwähnt dazu ein Pilotprojekt [Bro Jj, S. 3]. Zenodo erwähnt keine klassischen Statistiken. Eine Anzeige der Zugriffe oder Downloads auf Datensatzebene bieten Phaidra, LSERO, OceanRep und DR-NTU. OceanRep veröffentlicht außerdem eine Gesamtstatistik für die Nutzung der gesamten Webseite und MGDL zeigt nur eine Gesamtstatistik, aber wie erwähnt keine Zahlen auf Datensatzebene. Keines der FDR die Nutzungszahlen erheben trifft eine Aussage zur Erhebung und wie diese passiert. Lediglich DR-NTU verrät, dass die Metriken auf Daten aus Google Analytics und Piwik basieren [NTU17g].

K1.8 Anerkennung der wissenschaftlichen Leistung

Um Forscherx zu motivieren ihre Forschungsdaten zu veröffentlichen, werden in der Regel die erhöhte Sichtbarkeit (SDR, Zenodo, BDR, LSERO, OceanRep, DR-NTU), Anerkennung durch Zitate und erhöhte Impact Faktoren (Zenodo, BDR, OceanRep) genannt. LSERO und OceanRep werben mit Arbeitserleichterung in Form von automatisch erstellten und aktualisierten Publikationslisten. LSERO führt außerdem die Entstehung von Kollaborationsmöglichkeiten mit anderen Forscherx an [LSE14]. Phaidra nennt den Nutzerx keine Anreize. Eindeutige Autorenidentifikation ist in Zenodo, OceanRep und in Zukunft in DR-NTU über ORCID möglich. OceanRep nutzt außerdem auch die GND, ResearcherID und Scopus ID [Sch16]. Das SDR und LSERO haben keine solche Funktion und die übrigen drei geben keine Auskunft. Altmetrics für Forschungsdatensätze werden in Zenodo und LSERO angezeigt, SDR und DR-NTU befinden sich noch in der Planungsphase der Einbindung. MGDL stellt Datengeberx zweimal jährlich sog. "download & usage reports" [Mar Jd] zur Verfügung. Das BDR verweist auf externe Angebote wie ImpactStory und Google Scholar [Bro Jd]. Ein direktes Reporting an Förderer wird nur von Zenodo angeboten, über OpenAIRE kann der Europäischen Kommission Rechnung getragen werden [Zen Ja]. OceanRep ist in die interne Förderung und Mittelvergabe des tragenden Helmholtz Institutes eingebunden und liefert Zahlen für Jahresberichte und Evaluierungen [Sch14, S. 59].

6.2. K2 Zugriff und Auffindbarkeit

K2.1 Zugangsformen

Sieben der FDR gaben Hinweise zu den verschiedenen Zugangsbeschränkungen, die Forscherx für ihre Daten wählen können, darunter in jedem Fall die Veröffentlichung für alle Besucherx der Webseiten. Das SDR, BDR und OceanRep bieten eine Einschränkung auf Institutionsangehörige. Eine Beschränkung des Zugriffs auf einzelne Personen oder Nutzergruppen ist für Phaidra, Zenodo, BDR, OceanRep und DR-NTU möglich. Weiterhin erlauben Zenodo, BDR, OceanRep und MGDL einen "Closed access" oder privaten Zugang, in diesem Fall haben nur Datengeberx und Mitarbeiterx des FDR Zugang zu den Datensätzen. LSERO informiert hierzu nicht öffentlich. *Die Ergebnisse dieser Kategorie werden von der in Kapitel 2 beschriebenen Auswahl beeinflusst, wodurch alle FDR mindestens einen Datensatz öffentlich und einen zugriffbeschränkt angeben.*

Welche Nutzerx Daten in das FDR einspeisen dürfen wird überall angegeben. In sechs Fällen sind dazu nur Institutionsangehörige berechtigt, im LSERO erst ab Doktoranden-Status [LSE Je], im BDR Studierende nur wenn sie an Projekten teilnehmen [Bro Jf] und bei Phaidra auch externe Projektmitarbeiter [Abt Ji]. Zenodo und die MGDL nehmen FD von allen Personen entgegen, wobei in der MGDL überwiegend FD von Expeditionsteilnehmern eingespeist werden [Mar Je]. Eine Steuerung des Zugriffs für Daten, die nur bestimmten Personen zugänglich gemacht werden sollen (z.B. für Peer Review), ist für fünf Repositorien vorgesehen: Phaidra, Zenodo und das BDR bieten die Erstellung von Nutzergruppen an. Phaidra und OceanRep bieten einen begrenzten Zugriff und die MGDL einen passwortgeschützten Zugang [Mar Jd]. Dabei bleiben die Metadaten auch für zugangsbeschränkte Datensätze in Phaidra, Zenodo und OceanRep öffentlich sichtbar, die anderen FDR treffen dazu keine Aussage.

K2.2 Embargofunktion

Ein Embargo für FD festzulegen ist in sechs der Repositorien vorgesehen. Das BDR bietet Embargos für Dissertationen an [Cai17b], ob es für FD angewandt wird ist jedoch unbekannt. Nur das SDR gibt eine maximale Embargofrist an, diese beträgt drei Jahre [Sta Jd], für LSERO sind es mindestens zwei Jahre [Fla17]. Phaidra lässt die Sperrfristen vom Datengeberx festlegen [Abt Jf] und OceanRep orientiert sich an der Dauer die Forscherx extern auferlegt wird [GEO Jb]. Die übrigen vier informieren ihre Nutzerx dazu nicht. Die Anzeige einer Begründung für das Embargo konnte auf keiner der Webseiten gefunden werden, das Ende der Sperrfrist wird von Zenodo und der MGDL auf Datensatzebene angezeigt.

K2.3 Recherchemöglichkeiten

Sieben der FDR bieten eine Freitextsuche, davon in fünf Fällen getrennt in einfache und eine erweiterte Suche. Zenodo und das BDR haben nur eine Einschlitzsuche. Die native Suche der MGDL erlaubt keine freie Texteingabe, der übergeordnete IEDA Integrated Data Browser ermöglicht dies zwar extern, ist jedoch keine valide Datenquelle für diese Untersuchung. Das BDR und MGDL haben außerdem eine Search API, Zenodo testet eine Schnittstelle für die Suche und DR-NTUs Software ist dazu technisch fähig. Nur die Suchhilfe des SDR gibt eine Liste der möglichen Suchoperatoren an, OceanRep erwähnt nur Boolesche Operatoren [GEO Jb]. In Funktionstests war die Suche mit Booleschen Operatoren in vier weiteren Fällen erfolgreich (Phaidra, Zenodo, LSERO, DR-NTU). Die Ergebnisse für Suche mit Trunkierung war in den meisten ohne Fehlermeldung möglich, ebenso wie die Nutzung von Wildcards (ausgenommen OceanRep und DR-NTU), die Ergebnislisten waren jedoch nicht aussagekräftig genug um die Funktionalität zu bestätigen. Eine attributspezifische Suche ist für fünf der FDR vorgesehen, im MGDL sogar als ausschließlicher Einstieg [Mar Ji] als einfache, erweiterte und kartenbasierte Suche. Im BDR und in LSERO ist die Suche anhand ausgewählter Metadaten möglich.

Die Eingrenzung von Suchergebnissen über Facettierung ist für vier Repositorien möglich. Das BDR hat nur eine Auswahl der möglichen Facetten bereitgestellt und die Software von LSERO und OceanRep ermöglicht Facettierung ebensowenig wie die native Suche der MGDL. Empfehlungssysteme wurden für keines der betrachteten FDR implementiert. Volltextsuchen sind nur für textuelle FD relevant und in dieser Betrachtung nur im OceanRep implementiert, sowie potentiell in LSERO wie das Suchfeld "Documents" impliziert. Außerdem erwähnt das BDR ein Pilotprojekt für die Volltextsuche [Bro Jj, S. 1]. Die Landingpages der Datensätze sind in sechs Repositorien indexierbar und werden von externen Suchmaschinen gecrawlt, im BDR jedoch nur wenn öffentlicher Zugriff erlaubt wird [Bro Je]. Für die FDR die dazu keine Angabe machten wurde eine Known-Item-Suche durchgeführt, die im Fall von Zenodo erfolgreich war und für die MGDL nicht.

2.4 Integration in Zitationsdatenbanken

Nur zwei FDR erwähnen, dass MD an Zitationsdatenbanken oder -indexe geliefert werden: Zenodo verwendet u.a. DataCite [Zen Ja], OceanRep gibt an das für den Austausch gesorgt wird [GEO Ja], aber nicht wohin. Weiterhin haben sechs Repositorien eine Schnittstelle um den Austausch technisch zu ermöglichen, davon fünf über OAI-PMH. Zenodo hat zusätzlich dazu eine REST-Schnittstelle [Zen Jg]. Das MGDL hat mit dem MGDS FileServer eine eigene Schnittstelle [Mar Jh]. Erwähnenswert ist, dass Phaidra und DR-NTU ihre OAI Base-URL nicht auf der Webseite angeben, Phaidras ist jedoch über Verlinkungen zu erreichen.

K2.5 Programmierschnittstellen

Alle FDR haben APIs, überwiegend werden OAI-PMH Schnittstellen angeboten. Phaidra, Zenodo und die MGDL bieten auch eine REST API, während das SDR die Schnittstelle als RESTful bezeichnet¹⁷⁹. Das SDR hat außerdem eine IIIF Schnittstelle für Bilder [Sny16]. Das BDR hat drei APIs, gibt aber nicht an nach welchem Standard diese operieren [Bro Jc]. Die Funktionen dieser Schnittstellen reichen vom Upload (Phaidra, Zenodo), über die Suche (Zenodo, BDR, MGDL) bis zum Harvesting von MD, FD oder beidem. Im Fall des DR-NTU sind Schnittstellen wieder technisch möglich, zur praktischen Implementation liegen aber keine Hinweise vor [NTU17g].

6.3. K3 Transparenz und Qualitätssicherung

K3.1 Policies

Fünf der betrachteten FDR haben eine allgemeine Policy, das SDR hat nur eine Deposit Policy und LSERO nur eine Takedown Policy die auch explizit als Richtlinien benannt werden. Die weiteren Angaben die traditionell in Richtlinien erwartet werden, werden in allen Fällen über verteilte Hilfeseiten, Nutzungsbedingungen und FAQs präsentiert. Das SDR, LSERO und DR-NTU verweisen zudem für übergreifende Themen wie Open Access und FDM auf die Policies der Trägerinstitutionen.

K3.2 Datenvalidierung

Die Vollständigkeit der Metadaten wird im BDR, LSERO und in OceanRep geprüft. Zenodo und DR-NTU führen explizit keine Prüfung der Daten durch. Die übrigen drei Repositorien äußern sich dazu auf ihren Webseiten nicht. Zenodo, OceanRep und DR-NTU halten deutlich fest, dass die Verantwortung für den Inhalt der Daten bei den Datengeberx liegt, die anderen FDR lassen dies offen.

K3.3 Versionierung

¹⁷⁹Tiefergehende Ausführung zu REST / RESTful Bezeichnungen: <https://www.ibm.com/developerworks/webservices/library/ws-restful/>

Eine Versionierung von bereits veröffentlichten Datensätzen ist in sechs Repositorien möglich, nur LSERO und MGDL thematisieren dies Online nicht. Zur Kennzeichnung versionierter Datensätze bieten Phaidra, Zenodo und DR-NTU eine Versionsanzeige auf Datensatzebene. Im SDR steht nur die aktuelle Version online, weitere nur auf Anfrage [Sta Jd]. Phaidra und OceanRep verlinken alle Versionen miteinander. Die vorgenommenen Änderungen werden in drei FDR protokolliert und angezeigt: SDR, BDR und DR-NTU. Für die Vergabe der persistenten Identifizier hat jede Plattform eine eigene Lösung: In Phaidra behält die erste Version den PID, Zenodo vergibt einen PID pro Version und einen Übergeordneten pro Datensatz, das BDR zeigt alle Versionen unter einem PID und in OceanRep zeigt der PID auf die aktuellste Version. Die PID-Vergabe des DR-NTU ist nicht beschrieben.

K3.4 Löschung

Fünf Repositorien regeln die Löschung von Datensätzen auf ihren Webseiten, nur das SDR, LSERO und die MGDL informieren ihre Nutzerx dazu nicht. Die Angabe "Deleted Record Policy: persistent" [LSE Jg] in den LSERO OAI-PMH Daten deutet an, das eine Löschung möglich ist. Vier der Repositorien erhalten die Landingpages von gelöschten Datensätzen als Tombstone Pages, das sind Zenodo, OceanRep und DR-NTU, sowie wahrscheinlich LSERO. Ein Grund für die Löschung wird nur Zenodo Nutzerx angezeigt [Zen Jd]. Bei OceanRep muss ein Grund für die Löschung angegeben werden, ob dieser angezeigt wird konnte nicht ermittelt werden [GEO Jb].

K3.5 Sicherstellung langfristiger Zugänglichkeit

Alle analysierten Repositorien übernehmen auf unterschiedliche Weise die Verantwortung für den dauerhaften Zugang zu den FD. Das SDR definiert sich als "secure environment for long-term citation, access, and reuse" [Sta Jc], Phaidra verpflichtet sich "Datenbestände zu erhalten" [Ami J], Zenodo hat einen Abschnitt "Longevity" in den General Policies¹⁸⁰, das BDR wird als "Permanent, secure storage" beschrieben [Bro Jb], LSERO verspricht "future preservation and accessibility within the lifetime of the repository" [LSE Je], OceanRep hat einen Abschnitt "Preservation Policy" in den Repository Policies¹⁸¹, die MGDL zählt "long-term archiving" [Mar Jg] zu den Dienstleistungen und DR-NTU gibt an, dass alle FD "preserved and archived" [NTU17h] werden. Die angebotenen Sicherungsmaßnahmen sind ebenso vielfältig: Darunter ortsverteilte Aufbewahrung (SDR, Zenodo, BDR), redundante Datenhaltung (SDR, Zenodo, LSERO, DR-NTU), Bandsicherung (SDR, Zenodo), Format-, Datenträger und unspezifizierte Migration (Phaidra, BDR, DR-NTU) und die Überprüfung der Datenbestände (SDR, Zenodo, BDR), z.B. über Checksums. Die MGDL gibt ihre Sicherungsmaßnahmen nicht bekannt. Nur wenige FDR äußern sich zur genauen Sicherungsdauer. Bei Phaidra sind es mindestens 5 Jahre [Ami J], bei Zenodo mindestens 20 Jahre bzw. so lange der Träger CERN besteht [Zen Jd] und die Nutzer des DR-NTU sind durch die NTU Research Data Policy zu einer Aufbewahrung von 10 Jahren verpflichtet [Nan15], welches im Idealfall mindestens auf die Sicherungsdauer schließen lässt. Ebenfalls nur drei Repositorien äußern sich zur Nachfolge im Fall einer Schließung, davon benennt keines einen namentlichen Nachfolger. Phaidra plant eine Rückgabe der Daten an die Datengeberx, jedoch nur auf deren Anfrage [Ami J], Zenodo würde Daten an institutionelle oder disziplinäre Repositorien weiterleiten [Zen Jd] und OceanRep würde FD an ein Archiv abgeben [GEO Jc].

K3.6 Zertifizierung

Keines der betrachteten Repositorien ist zertifiziert.

¹⁸⁰ <http://about.zenodo.org/policies/>

¹⁸¹ <http://oceanrep.geomar.de/policies.html>

6.4. K4 Nutzererlebnis und Support

K4.1 Usability

Um ihre Nutzerfreundlichkeit zu erhöhen haben vier FDR besondere Maßnahmen ergriffen. Phaidra bietet die Benutzeroberfläche in Deutsch, Englisch und Italienisch an [Abt Jf] und lässt Forscherx Vorlagen für häufig eingegeben Metadaten anlegen [Abt08, S. 25]. LSERO wirbt mit einer Guideline zum Datenexport vom FDR nach ORCID [LSE Jc] und bietet automatisch erstellte und aktualisierte Publikationslisten für alle Einträge im Repository an [LSE Je]. OceanRep stellt ebenfalls Publikationslisten bereit [Sch14], ermöglicht die Erstellung von Vorlagen im Ingest und weist auf unvollständige Eingaben im Upload- Formular hin [GEO Jb]. Die MGDL bietet eine Alternative zur ressourcenintensiven Kartensuche für langsame Internetverbindungen (siehe Screenshot MGD010_Kartensuche) und zeigt die neuesten Einträge für aktuelle Forschungsexpeditionen¹⁸² an. Zenodo verschlechtert die Usability mit unspezifischen Fehlermeldungen (siehe Screenshot Zen015_Usability-Error). SDR und DR-NTU stechen weder positiv noch negativ heraus.

Die Bedienung schätzt die Autorin für fünf der Repositorien als intuitiv ein, die Suchfunktionen auf den Webseiten verhalten sich entsprechend den Erwartungen. Bspw. zeigt eine Suche mit leeren Feldern alle Ergebnisse an, eine Anpassung der Suche erfolgt ohne Seitenwechsel, eine Freitextsuche ist möglich, Suchparameter bleiben bei Rückkehr auf vorige Seiten bestehen, etc.. Die auf EPrints basierenden Entwicklungen LSERO und OceanRep fallen negativ auf, da sie z.B. keine leere Suche erlauben (siehe Screenshot LSE010_LeereSuche) und die Suchanfragen nicht in der Anzeige der Ergebnisse angepasst werden können (siehe Screenshot ORE014_SucheAnpassen). Die Bedienung der kartenbasierten Suche in der MGDL ist für Fachfremde nicht ausreichend erklärt (siehe Screenshot MGD010_Kartensuche) und die Suche über Metadaten ist ungewohnt. Die Darstellung der Symbole entspricht auf allen Webseiten den gängigen Formen, wenn auch LSERO und OceanRep erneut negativ auffallen, da die in EPrints verwendeten Icons veraltet wirken. Ein Beispiel ist das Lupensymbol des LSERO (siehe Screenshot LSE001_AnzeigeDatensatz) im Vergleich zu dem, dass im neu gegründeten DR-NTU (siehe Screenshot NTU002_Verlinkung) verwendet wird. Die im Idealfall flache Lernkurve wird für sechs der FDR durch eine fehlende Suchhilfe beeinträchtigt, wobei die MGDL eine eingeschränkte Hilfeseite bietet. Nutzerx müssen z.B. ausprobieren welche Suchoperatoren angewandt werden können um eine präzise Suche durchzuführen. Das SDR und OceanRep unterstützen ihre Nutzerx mit umfangreichen Suchhilfen.

Keines der betrachteten Repositorien erreicht die niedrigste Konformitätsstufe A nach den Anforderungen der WCAG 2.0¹⁸³ [Cal+09], d.h. beispielsweise das alle gegen Richtlinie 1.1 "Textalternativen" verstoßen und keine Textalternativen für Bilder, Videos u.ä. anbieten, entgegen Richtlinie 1.2 "Zeitbasierte Medien" keine Untertitel oder Transkriptionen für Videos anbieten, die Darstellung der Webseiten nicht wie in Richtlinie 1.3 "Anpassbar" gefordert den Sehfähigkeiten von Nutzerx angepasst werden kann usw. Nur das BDR kann entsprechend der Richtlinie "2.1 Per Tastatur zugänglich" ohne Maus benutzt werden. Die Browsererweiterung aXe prüft maschinenlesbare Verstöße gegen die WCAG (darunter Richtlinie 1.4 "Unterscheidbar", Richtlinie 3.3 "Hilfestellung bei der Eingabe" und Richtlinie 4.1 "Kompatibel") und fand auf allen Webseiten diverse Verstöße. Geprüft wurde in der Anzeige der Suchergebnisse und in der Anzeige eines Datensatzes. Die Suchergebnisse zeigten zwischen 11 und 173 Fehler und die Datensatzansicht zwischen 3 und 46 Verstößen. Die niedrigste Anzahl an Verstößen ins-

¹⁸²<http://www.marine-geo.org/portals/antarctic/news.php>

¹⁸³ Deutsche Übersetzung der Web Content Accessibility Guidelines: <https://www.w3.org/Translations/WCAG20-de/>

gesamt hatte LSERO, gefolgt von Phaidra. Die meisten Verstöße wurden auf den Webseiten von Zenodo und DR-NTU angezeigt. Eine Aufzählung der häufigsten Verstöße pro FDR ist in Anhang A.2 unter Punkt K4.1 Usability - Barrierefreiheit verfügbar, es handelte sich überwiegend um zu niedrigen Farbkontrast oder fehlende Beschriftungen.

Vier der analysierten FDR zeigen Bemühungen: Phaidra bietet Suchvorschläge an (siehe Screenshot PHA014_Accessibility-Suchvorschläge). Die Brown University hat einen Closed Captioning Service für Videos [Bro Ja] und die IT-Abteilung plant eine Accessibility Prüfung für alle betreuten Projekte [Bro Jj]. Die LSE nennt ihr Web Services Team als Ansprechpartner für Barrierefreiheit auf allen Universitätswebseiten [Lon J]. Und die MGDL hat eine Auto-Vervollständigung in ihrer Suchfunktion [Mar Jh].

K4.2 Darstellung der Datensätze

Datenvisualisierung wird von Phaidra für un spezifizierte Ressourcen angeboten [Abt Jb], das BDR bietet Beratung zur Visualisierung von Datensätzen an [Bro Jg], die MGDL bietet mehrere Visualisierungen für geografische und zeitliche Daten und div. 3D Browser [Mar Jg] und DR-NTU bietet Werkzeuge für die Visualisierung von geografischen und tabellarische Daten ([NTU17b, S. 1 u. 3]). Eine Vorschaufunktion für Videos und Bilder bieten das SDR, Phaidra, das BDR, LSERO und OceanRep, außerdem haben Zenodo und die MGDL eine Bildvorschau. Im DR-NTU Repository sind bisher keine Bilder hochgeladen, Videos haben jedoch keine Vorschau- oder Streamingoption.

K4.3 Support zu Repositoriums-funktionen

Sieben der analysierten FDR haben Hilfeseiten oder FAQs um Nutzerx über die Funktionen des Repositoriums zu informieren. Phaidra hat kein explizit als Hilfeseite betiteltes Angebot, unterhält aber diverse Seiten zu spezifischen Themen und DR-NTU nutzt die Dataverse User Guidelines nach (dazu mehr im DR-NTU Kriterienkatalog in Anhang A.2). Nutzerxschulungen werden in unterschiedlichen Formen vom SDR, Phaidra, dem BDR und der MGDL durchgeführt, bei DR-NTU sind diese noch in Planung [Wen17]. Zenodo, LSERO und OceanRep bieten keine Schulungen zur Benutzung des FDR auf den Webseiten an. Supportanfragen können in vier Repositorien über ein Kontaktformular gestellt werden (SDR, Zenodo, MDGL, DR-NTU), das Nutzerx informiert welche Angaben für die Behebung von Fehlern hilfreich sind. Die übrigen vier FDR stellen eine E-Mail Adresse zur Verfügung. Allein Nutzerx des MGDL werden explizit gebeten inhaltliche Fehler in den FD oder MD an die Mitarbeiterx zu melden [Mar Jj], kein weiteres analysiertes Repository bietet dazu die Möglichkeit.

K4.4 Beratung zum Forschungsdatenmanagement

Angebote die durch eine direkt angeschlossene Bibliothek angeboten werden, wurden mit einbezogen da Online- Angebote auch für externe Nutzerx unmittelbar erreichbar und Offline-Angebote mindestens für Institutsangehörige Forscherx verfügbar sind. Sechs der Repositorien beraten zum übergeordneten Thema des FDM, lediglich Zenodo und die MGDL bieten auf ihren Webseiten keine Angebote. Konkrete Hilfestellungen bis hin zur Übernahme einzelner Arbeiten des FDM bieten dieselben Repositorien, das SDR z.B. Datenmanagement und Digitalisierung [Sta Jd], Phaidra begleitet Projekte ab Antragsstellung [BG17, S. 23] und hat die Phaidra-Sandkiste als Testumgebung [Abt Je], das BDR und LSERO speisen Daten für ihre Nutzerx ein usw.. Ebenso verweisen alle sechs auf aktuelle Richtlinien und Standards des FDM: LSERO verweist auf SHERPA/RoMEO¹⁸⁴ [LSE Je] oder die UK Data Service Recommended Formats [LSE Jb], OceanRep empfiehlt Forschungsdaten.org [GEO15] und DR-NTU nutzt die Harvard Dataverse Policies nach und verweist auf re3data [NTU17n].

¹⁸⁴ <http://www.sherpa.ac.uk/romeo/index.php>

6.5. K5 Organisation und Infrastruktur

K5.1 Organisationsform

Sechs der Repositorien sind in einer Abteilung der Bibliothek des Trägers angesiedelt. Im Fall von Phaidra und DR-NTU erfolgt außerdem eine Betreuung durch die IT-Abteilung der Universität. Zenodo wird nur durch die IT-Abteilung des Trägers CERN geleitet und das MGDL ist in ein Forschungsinstitut der Columbia University eingebettet. Bei OceanRep kooperiert die Bibliothek mit dem Datenmanagementteam des tragenden Helmholtz-Institutes. Zur Langfristigkeit der Organisationsform oder des Trägers macht nur Zenodo eine Aussage: CERN wird mindestens 20 Jahre bestehen und Zenodo ist direkt an diese Laufzeit gebunden [Zen Jd]. Alle analysierten FDR sind an etablierten Hochschulen oder Instituten angesiedelt. Die Träger werden auf allen untersuchten Webseiten deutlich genannt, es sind hauptsächlich Universitäten. Träger von Zenodo sind die Europäische Kommission und CERN [Zen Je], für OceanRep die im Helmholtz-Zentrum angesiedelten Sonderforschungsbereiche und der Exzellenz-Cluster [Sch14, S. 60] und für die MGDL die US-Amerikanische National Science Foundation [Mar Ja]. Zur Gemeinnützigkeit des Repositoriums gibt nur das SDR eine Aussage und bestätigt dies [Sta JI]. Bei OceanRep wird erwähnt, dass der Träger gemeinnützig ist [Sch14, S. 58]. Die sechs anderen FDR informieren ihre Nutzerx nicht pro-aktiv.

K5.2 Personalausstattung und -qualifizierung

Die Anzahl der Mitarbeiter, die an den Repositoren beteiligt sind wird selten deutlich genannt, deren Qualifikationen bleiben auch oft offen. Das SDR beschäftigt 30 Personen, darunter Experten für Dienstleistung, LZA und IT [Sta Ji]. Die Universität Wien zeigt 3 Mitarbeiter für die Abteilung Phaidra-UNIDAM¹⁸⁵ an, sagt aber nichts zu deren Qualifikation aus. CERN bezeichnet das Personal als Experten und zeigte im Dezember 2017 für die gesamte Abteilung 15 Mitarbeiter¹⁸⁶ an, erwähnt aber auch, dass nur ein Teil für Zenodo zuständig ist [Zen Je]. Die Brown University informiert nicht wer am BDR mitarbeitet, die Abteilung "Center for Digital Scholarship" hat aber 8 Mitarbeiter¹⁸⁷ inkl. einem Digital Repository Manager und einem Data Visualization Coordinator und die Abteilung Digital Technologies hat 3 Mitarbeiter für das BDR¹⁸⁸: Je ein Digital Repository Manager, Digital Library Programmer und Digital Preservation Librarian. LSERO wird von vier bibliothekarischen Mitarbeiterx betreut, die gleichzeitig auch andere Ressourcen katalogisieren [Wil17]. Die Bibliothek, die das OceanRep betreibt hat 3¹⁸⁹ und das Kieler Datenmanagement-Team hat 5 Mitarbeiter¹⁹⁰, es fehlen jedoch Hinweise wer an OceanRep mitarbeitet und in welcher Kapazität. Die MGDL wird von 9 Mitarbeiterx und 2 Projektleiterx betreut [Lam J], dabei handelt es sich nach allen Angaben um Fachwissenschaftlerx. Das "Research Data Management Team" der NTU Bibliothek verzeichnet 3 Mitarbeiterx [NTU Jb], gibt aber keine Informationen zu deren Qualifizierung.

K5.3 Forschungs- und Entwicklungstätigkeiten

Hinsichtlich der Weiterentwicklung der Infrastruktur wurden die Webseiten der zuständigen Abteilungen oder des Trägers analysiert. Die Stanford University entwickelt das Repository-Framework "Hydra-in-a-Box", welches auch Grundlage des SDR ist [Sta17]. Phaidra ist eine Eigenentwicklung der Universität Wien und wird dort weiterentwickelt [Abt Jf], außerdem nimmt es an den Infrastrukturprojekten OpenAIRE und e-Infrastructures Austria teil [Abt Ja]. Zenodo ist Projektpartner von OpenAIRE und EUDAT [Zen Je], gibt aber keine Details zur

¹⁸⁵<http://datamanagement.univie.ac.at/ueber-uns/>

¹⁸⁶<https://phonebook.cern.ch/phonebook/#search/?query=IT-CDA-DR>

¹⁸⁷<https://library.brown.edu/cds/about/people/>

¹⁸⁸<https://library.brown.edu/dt/>

¹⁸⁹<https://www.geomar.de/zentrum/einrichtungen/bibliothek/team/>

¹⁹⁰<https://www.geomar.de/de/zentrum/einrichtungen/rz/daten/>

Beteiligung. Die Brown University Abteilung CDS nennt die Weiterentwicklung von Repositorien als Kernziel [Bro Jh]. Die Datenbank auf der die MGDL basiert, wurde intern entwickelt [Mar Jf]. Die Unterstützung von allgemeinen Projekten wird nicht thematisiert. Die Webseiten von LSERO, OceanRep und DR-NTU geben keine Hinweise auf Infrastrukturentwicklungen. Im Bereich des Forschungsdatenmanagements oder der betreuten Fachdisziplin bemühen sich vier der untersuchten Repositorien: Die Stanford Universitätsbibliothek hat ein Plugin für die Erstellung von Webseiten für digitale Sammlungen entwickelt [Sny14]. Phaidra ist involviert im Aufbau des "Netzwerk für RepositorienmanagerInnen" in Österreich [BF J] Und die zuständige Abteilung des BDR führte bereits diverse Digitalisierungsprojekte mit der Fachdisziplin¹⁹¹ durch.

Weiterhin arbeiten drei Repositorien an der Weiterentwicklung von FDM-Standards: Die Stanford University Library kooperierte mit anderen Organisationen um das "Manifesto for Linked Libraries"¹⁹² zu entwickeln. Zenodo ist einer der Unterzeichner der "Joint Declaration of Data Citation Principles"¹⁹³. Und die Mitarbeiter der Abteilung, die auch das BDR betreuen, geben an, eine aktive Rolle in der Entwicklung von neuen Metadatenstandards einzunehmen [Bro Jh]. Weitere Details werden jedoch nicht genannt.

Außerdem listen sechs der Repositorien diverse nationale, internationale und fachspezifische Forschungsprojekte auf ihren Webseiten auf. Nur LSERO, die MGDL und DR-NTU informieren ihre Nutzerx dazu nicht öffentlich.

K5.4 Technische Infrastruktur

Alle Webseiten nennen mindestens allgemeine technischen Grundlagen für ihre Repositorien. Das SDR wird auf dem Hydra Framework betrieben, dieses beinhaltet Fedora (Basis des Repositoriums), Solr (Suche), Blacklight (Facettierung und Browsing) und Moab für die Versionierung [Sta Jj]. Phaidra basiert ebenfalls auf Fedora und läuft unter dem Betriebssystem Linux [Gan17, S. 110]. Zenodo läuft auf dem Invenio Framework [Zen Je]. Das BDR nutzt das veraltete Fedora 3 und sucht nach einem neuen Framework [Cai17a]. LSERO und OceanRep nutzen EPrints in der Version 3.3 ([LSE Je] und [Sch14, S. 57]). Die MGDL basiert auf einer relationalen Datenbank, die für das FDR entwickelt wurde [Mar Jf]. Und DR-NTU ist eine Instanz von Dataverse [NTU17a].

Aussagen zu den Ladezeiten sind subjektiv und basieren auf den Erfahrungen der Autorin in der Erstellung der Analyse. Die einzigen auffälligen Abweichungen wurden in der Suche in LSERO, OceanRep und der Kartensuche der MGDL festgestellt, die wesentlich langsamer waren als erwartet. Zenodo hat außerdem im September 2017 die Geschwindigkeit der Suche durch die Einbindung von Elasticsearch optimiert [Zen17].

Nur zwei Repositorien machen Aussagen zu der Ausfallsicherheit ihrer Systeme: Zenodo läuft auf zwei getrennten Systemen um alle Updates vor Release testen zu können ohne dass Nutzerx es merken oder von Ausfällen betroffen werden [Zen Je] und das BDR betreibt vier separate Server für die diversen Services, darunter ein Server allein für die Suchfunktion [Cai17c].

Zur Backup-Strategie äußern sich drei der Repositorien: Das SDR sichert redundant auf Festplatten und Magnetband, gibt aber nicht an wie oft Kopien gemacht werden [Sta Jj]. Zenodo sichert FD nächtlich auf Band und MD alle 12 Stunden [Zen Je]. Und OceanRep gibt nur an regelmäßige Backups zu ziehen und sich dabei an Best Practice zu halten, gibt aber keine Details wie oft und auf welche Datenträger die FD überspielt werden [GEO Jc].

Nur Zenodo gibt an eine Überwachung der Hardwarekomponenten durchzuführen und nutzt dafür die allgemeine CERN Infrastruktur mit, außerdem werden sämtliche Fehlermeldungen

¹⁹¹ <https://library.brown.edu/cds/projects/>

¹⁹² <https://library.stanford.edu/projects/linked-data>

¹⁹³ <https://www.force11.org/datacitation/endorsements>

aggregiert [Zen Je].

Eine Dokumentation der technischen Infrastruktur ist für Phaidra Online in github¹⁹⁴ erhältlich, ebenso wie die Dokumentation der phaidra-API¹⁹⁵. Zenodo hat den gesamten SourceCode ebenfalls auf github¹⁹⁶ dokumentiert, sowie die entwickelten APIs auf dem Developers blog¹⁹⁷. Alle Repositorien die technische Informationen bereitstellen, nennen überwiegend offene Software. Lediglich die Angaben der MGDL waren nicht ausreichend um dies einzuschätzen.

6.6. K6 Gebühren und Finanzierung

K6.1 Gebühren für die Datenabgabe und -nutzung

Vier der Repositorien geben explizit an, dass das Angebot kostenfrei genutzt werden kann. Zenodo akzeptiert zusätzlich Spenden und erwähnt dies im Kontext des Uploads großer Datensätze [Zen Jc]. Das SDR entwickelt derzeit eine Gebührenmodell [Sta Jg]. Für die übrigen drei ist anhand der Analyse deutlich, dass die Nachnutzung von Daten kostenfrei ist. In keinem Fall wurde ein Vertrag zwischen den Institutionen oder deren Abteilungen und dem FDR erwähnt, der die Finanzierung regelt oder Gebühren festlegt.

K6.2 Finanzierung des Betriebs

Zenodos Webseite erwähnt, dass das FDR an die Laufzeit von CERN gebunden ist, welches ein Konzept für die nächsten 20 Jahre angibt [Zen Jd]. Ob und wie Zenodo darin enthalten ist bleibt unklar. Der "Brown University Libraries Technology Plan, 2015-2017"¹⁹⁸ enthielt Pläne für das BDR, bisher liegt kein Hinweis auf einen Plan ab 2018 vor. OceanRep wird vom Exzellenz-Cluster 'Future Ocean' und den Sonderforschungsbereichen 574 und 754 mitfinanziert, die ebenfalls im Helmholtz- Institut ansässig sind ([Sch14, S. 60]). Dies deutet auf eine langfristige Finanzierung hin; ein Konzept ist jedoch nicht öffentlich verfügbar. OceanRep ist das einzige analysierte Repository, das eine Aussage zum Budget trifft. Hierbei handelt es sich um eine Schätzung auf ca. 3000€/Jahr, die 2014 im Rahmen eines Plattformwechsels geäußert wurde (ebd.). Keines der FDR trifft öffentlich Prognosen für die Finanzierung in der Zukunft.

¹⁹⁴<https://github.com/phaidra>

¹⁹⁵<https://github.com/phaidra/phaidra-api/wiki/Documentation>

¹⁹⁶<https://github.com/zenodo/zenodo>

¹⁹⁷<http://developers.zenodo.org>

¹⁹⁸<https://library.brown.edu/dt/LORITechnologyPlan2015.pdf>

7 Bewertung der Ergebnisse

In dieser Arbeit wurden acht Repositorien untersucht, die audiovisuelle Forschungsdaten vorhalten. Ausgewählt wurden die Repositorien aus einer Grundgesamtheit die über das Verzeichnis re3data ermittelt wurde, dabei ist zu beachten das re3data nicht alle existierenden Repositorien verzeichnet. Die Auswahl der näher zu betrachtenden Repositorien geschah anhand festgelegter Kriterien aus dem re3data Metadatenchema und funktionalen Anforderungen an die Suchoptionen, deren Einhaltung in einer ersten Analyse überprüft wurden. Die Grundlage dieser Arbeit ist eine inhaltliche Analyse der Webauftritte der Repositorien anhand des Kriterienkatalogs von Steinhoff [Ste17]. Die Angaben der Repositorien für jede Kategorie wurden miteinander verglichen. Dies erlaubt eine exemplarische Betrachtung von Gemeinsamkeiten und Unterschieden für die betrachteten Aspekte zwischen den analysierten Repositorien.

Grundsätzlich wurde die Auswahl der Repositorien in re3data durch die Wahl des Metadatum "dataAccessRestrictions" mit dem Wert "registration" beeinträchtigt. Die in der ersten Analyse (siehe Anhang A.1) untersuchte Anzahl an Repositorien ist eine Teilmenge der Repositorien die ursprünglich untersucht werden sollten. Tatsächlich untersucht wurden nur Repositorien die mindestens einen Datensatz im Zugriff beschränken, dennoch waren alle Repositorien durch das Metadatum "databaseAccessType" mit dem Wert "open" Teil der Menge die untersucht werden sollte. Die vorliegende Untersuchung gibt also keine Hinweise auf FDR deren Daten vollständig offengelegt werden und die daher in "dataAccessRestrictions" keinen Wert angegeben haben. In der Analyse der 92 Repositorien aus re3data (siehe Anhang A.1) erfüllten nur 13 Repositorien (knapp 12%) die festgelegten Kriterien. Dies beinhaltet eine Abweichung von den in re3data hinterlegten Angaben (contentType, dataAccess, repositoryLanguages) in 15% der Fälle (17 von 92) und entsprach nicht den Erwartungen die in der redaktionellen Arbeit zu Erstellung von Einträgen re3data begründet lagen [Pam14, Folie 17]. Weiterhin wurde deutlich, dass keines der kriterienkonformen Repositorien einen hohen Anteil von Videodaten vorwies. Der höchste erreichte Wert lag bei 25% Videos gegenüber der Gesamtanzahl an Datensätzen in der National Science Digital Library. Der Ausschluss von Repositorien die ihre Daten aus anderen Quellen aggregieren und nicht durch Personen einpflegen lassen führte dazu, dass keines der analysierten Repositorien mehr als 2% Videos enthält. Der niedrigste prozentuale Anteil von Videos in einem untersuchten FDR liegt bei 0.03% im MGD, die niedrigste totale Anzahl von Videos liegt bei 1 im DR-NTU.

Wie erwartet gab kein Repository Hinweise zu allen Kriterien. Der Vergleich der Repositorien miteinander zeigte aber eine Reihe von Gemeinsamkeiten und Unterschieden im Umgang mit den untersuchten Kriterien auf; vor allem waren aber Tendenzen zu erkennen.

K1 Veröffentlichung und Nachnutzung

Die betrachteten FDR tendieren dazu alle Dateiformate zu akzeptieren, empfehlen aber gleichzeitig eine Reihe an archivfreundlichen Formaten. Der Upload von Forschungsdaten erfolgt meist über den Browser, alternativ werden Schnittstellen oder eine Ingest-Dienstleistung angeboten. Laut den Webseiten, wird die Integrität der Daten überwiegend nicht geprüft. Die Repositorien die Daten überprüfen tun dies intellektuell. Dies widerspricht den Angaben in re3data, denn dort sind alle FDR als "qualityManagement:yes" geführt (z.B. [re317f]), entsprechend der Auswahl der Repositorien die in Kapitel 2 beschrieben wurde. Die Webseiten treffen jedoch keine Aussage zu Qualitätsmanagement oder verneinen dies sogar.

Es gab keinen Konsens zur Struktur des Metadatenchemas (auch "Metadatenkategorien")

genannt). Angaben dazu wurden nur von weniger als der Hälfte der Repositorien gemacht. Eine deutlichere Trennung in beschreibende, technische und administrative Metadaten wurde aufgrund der Fachliteratur (z.B. [RS16, S. 26], [nes16, S. 55]) erwartet. Keines der Repositorien nutzt einen einzelnen, unveränderten Metadatenstandard. Standards wurden miteinander kombiniert und an die Bedürfnisse angepasst. Keines der FDR hat sein vollständiges Metadatenchema auf der eigenen Webseite veröffentlicht.

Nutzer werden auf vielfältige Weise bei der Eingabe von Metadaten unterstützt, eine Tendenz bildet sich nicht ab. Mehr als einmal genannt wurden Autovervollständigung von Feldern im Eingabeformular und eine Speichermöglichkeit für Vorlagen.

Alle Repositorien die dazu Auskunft geben bieten einen formularbasierten Upload, in dem auf Pflichtangaben hingewiesen wird.

Tendenziell werden kontrollierte Vokabulare für die Vergabe von Schlagworten angeboten, deren Form ist aber unterschiedlich und reicht von nationalen Thesauri bis zu Fachontologien.

Eine Dokumentation der Forschungsdaten ist laut den Webseiten in keinem der FDR, die dazu eine Angabe machen verpflichtend. Es ist Nutzerx aber möglich diese als Datei zu inkludieren oder extern zu verlinken.

Explizite Zitationsrichtlinien wurden selten formuliert; ein Beispiel für eine ausführliche Richtlinie bietet der UK DataService¹⁹⁹. Stattdessen werden überwiegend Zitationsvorschläge angezeigt an denen sich nachnutzende Personen orientieren können. Standards wie die Data Citation Principles werden in den betrachteten Repositorien selten eingesetzt.

Alle FDR geben auf ihren Webseiten Informationen zur dauerhaften Zitierfähigkeit der vorgehaltenen Daten. Eine deutliche Tendenz besteht in der Nutzung von DOIs, andere Systeme wurden nur in Einzelfällen genannt.

Die Festlegung von Nutzungsbedingungen für Daten erfolgt sehr heterogen, mehrfach wurden jedoch allgemeine Richtlinien oder die Terms of Use der Webseite als Ort gewählt. Deutlich gemein sind allen untersuchten Repositorien die Möglichkeit Daten offen zu lizenzieren, dies beinhaltet für alle mindestens die Creative Commons Lizenzen und in Einzelfällen weitere Standardlizenzen. Die Metadaten werden von den Repositorien eher nicht explizit lizenziert. Die Webseiten, die Angaben machen, stellen ihre Metadaten bindungslos zur Verfügung oder untersagen kommerzielle Nutzung.

Eine Verknüpfung von Forschungsdaten und daraus resultierenden Publikationen ist fast überall explizit vorgesehen und technisch immer möglich. Die automatische Verlinkung von FD und Publikationen, die diese nachnutzen, ist im Gegensatz dazu auf keiner Webseite erwähnt, manuell möglich ist es aber in einigen Repositorien über Metadatenfelder die auf verwandte Inhalte zeigen. Alle analysierten Repositorien sehen die Verknüpfung mit zugehörigen Projekten oder anderen Forschungsdaten vor, welches eine weitere komplette Gemeinsamkeit darstellt. Unterschiedlich ist die Umsetzung dieser Verknüpfung, mehrfach genannt wurden aber digitale Sammlungen.

Nutzungsstatistiken werden von einer knappen Mehrheit erhoben, meist werden die Zugriffs- oder Downloadzahlen auf den Landingpages angezeigt. Keines der FDR, die Nutzungszahlen erheben trifft eine Aussage zur Erhebung der Daten oder nutzt eine standardisierte Erhebung wie den COUNTER Code of Practice²⁰⁰.

Als Anreiz für die Veröffentlichung von Forschungsdaten werden tendenziell eher erhöhte Sichtbarkeit und Anerkennung genannt; ein Konsens ist nicht zu erkennen. Wenn Systeme für die eindeutige Autorenidentifikation genutzt werden liegt die ORCID in den untersuchten Reposi-

¹⁹⁹ <https://www.ukdataservice.ac.uk/use-data/citing-data>

²⁰⁰ Mehr Informationen zu COUNTER: <https://www.projectcounter.org/code-of-practice-sections/general-information/>

torien vorn, die Mehrheit thematisiert dies aber nicht oder bietet keinen entsprechenden Service an. Altmetrics werden nur selten eingebunden, sind jedoch in zwei Repositorien bereits in Planung. Direktes Reporting des Forschungsausgangs an Förderer wird nur in einem Fall angeboten. Keines der institutionellen Repositorien weist auf eine solche Hilfestellung für ihre Nutzerx hin. In einem weiteren Fall werden extrahierte Listen aus dem FDR für interne Berichte genutzt.

K2 Zugriff und Auffindbarkeit

Die deutliche Mehrheit der Repositorien informiert auf ihrer Webseite welche Zugangsbeschränkungen festgelegt werden können. Die Ergebnisse dieser Kategorie werden von der in Kapitel 2 beschriebenen Auswahl beeinflusst: FDR die alle Datensätze öffentlich anbieten wurden ausgeschlossen. Daher bieten alle untersuchten FDR ihren Nutzerx die Möglichkeit ihre Datensätze nicht- öffentlich zu hinterlegen. Grundlegend folgen die FDR den vier Arten von Zugangsbeschränkungen, die von Rice und Southall vorgeschlagen wurden: Open, Managed, Secure und Closed [RS16, S. 111]. Ein FDR gab keine Hinweise zu diesem Aspekt, die anderen bieten alle mindestens die freie Veröffentlichung ("Open") und eine Einschränkung. Deren Granularität ist unterschiedlich, tendiert aber zu Zugang nur für Institutionsangehörige, einzelne Nutzergruppen oder Einzelpersonen ("Managed"). Auch privater Zugang nur für den Datengeber ("Closed") wird mehrfach angeboten. Sicherer Zugang für sensible oder geheime Daten ("Secure") wird nicht angeboten.

Alle Analysen fanden Angaben zum Nutzerkreis der zur Datenabgabe berechtigt ist. In den institutionellen Repositorien waren dies mehrheitlich Institutionsangehörige, teilweise mit Einschränkungen für Studierende. Sensible oder unveröffentlichte Daten können in über der Hälfte über Nutzergruppen oder Passwortschutz zugänglich gemacht werden. Ob die Metadaten dabei öffentlich bleiben geben nur wenige Repositorien an, bei diesen ist es aber der Fall.

Viele FDR erlauben ein Embargo, eine feste Frist wird jedoch selten genannt und noch seltener dem Nutzerx angezeigt. Die Anzeige einer Begründung für das gesetzte Embargo konnte in der Analyse an keiner Stelle gefunden werden.

Überwiegend ist die Suche nach Forschungsdaten über Freitextfelder möglich. Über die Hälfte der Repositorien folgen der klassischen Trennung in einfache und erweiterte Suche. Schnittstellen für die Suche werden in zwei Fällen angeboten und sind in einem weiteren in der Testphase. Suchhilfen werden nur selten angeboten und beinhalten nur in einem Fall die Suchoperatoren. In einem Fall ist die Suche ausschließlich attributspezifisch vorgesehen, für eine Überzahl der Repositorien ist sie ebenfalls möglich. Eine Facettierung ist in der Hälfte der FDR implementiert. Obwohl mindestens ein FDR in ein Discovery System eingebettet ist und alle FD erschlossen sind, wurde in dieser Untersuchung kein Repository mit Empfehlungssystem gefunden. Volltextsuchen sind nur selten möglich, jedoch für Forschungsdaten im Allgemeinen und Audiovisuelle im Besonderen weniger relevant. Eine Tendenz zeigt sich bei der Indexierbarkeit der Landingpages. Die meisten Repositorien geben an, dass ihre Webseiten gecrawlt werden und die Forschungsdaten somit in Suchmaschinen zu finden sind.

Die Lieferung von Metadaten an Zitationsdatenbanken wie DataCite wurde nur in zwei Fällen thematisiert. Die Mehrheit der Repositorien hat aber eine Metadaten- Schnittstelle. Meist wird OAI-PMH genutzt.

Gemeinsam ist allen Repositorien, dass sie eine API anbieten, darunter OAI-PMH, REST und Eigenentwicklungen. Diese Schnittstellen werden aber für unterschiedliche Funktionen angeboten u.a. für den Upload von Datensätzen, Suche im Datenbestand oder Harvesting.

K3 Transparenz und Qualitätssicherung

Allgemeine Policies konnten in über der Hälfte der Repositorien festgestellt werden, alle Webseiten hatten jedoch gemein dass Informationen auf verschiedene Hilfeseiten, in den Nutzungsbedingungen oder FAQs verteilt wurden. In vielen Fällen mussten auch die Seiten der Träger

einbezogen werden und in Ausnahmen auch Publikationen der Mitarbeiterx. Dies ist für Nutzerx nicht transparent.

Die Vollständigkeit der Metadaten wird eher selten überprüft und in den untersuchten Repositorien vor allem manuell durch Mitarbeiterx. In zwei Fällen wird sogar explizit keine Prüfung vorgenommen. Die Verantwortung für den hochgeladenen Inhalt wird selten explizit geklärt, die Repositorien die es tun benennen die Datengeberx.

Die Repositorien erlauben überwiegend die Versionierung von bereits veröffentlichten Datensätzen, aber nicht alle nehmen dazu Stellung. Die Kennzeichnung der Versionierung geschieht oft über eine Versionsanzeige, weitere Gemeinsamkeiten finden sich jedoch nicht. Vor allem die Vergabe von PID für versionierte Datensätze ist für jedes FDR anders geregelt. Ein Protokoll der Änderungen wird in weniger als der Hälfte der Fälle erstellt.

Eine Löschung von Datensätzen ist in knapp über der Hälfte der Repositorien geregelt, in vier Fällen ist es nur in Ausnahmefällen erlaubt. Ein FDR erlaubt die Löschung von Datensätzen bei Ausscheiden des Datengebers aus dem Institut, welches nicht den Grundsätzen der dauerhaften Sicherung entspricht. Die Landingpages der FDR die dazu Angaben machen bleiben erhalten. Die Hälfte der Repositorien thematisiert den Vorgang der Löschung aber nicht. Nur in einem Fall wird ausgesagt, dass der Grund der Löschung angezeigt wird.

Alle analysierten Repositorien übernehmen auf unterschiedliche Weise die Verantwortung für den dauerhaften Zugang zu den FD. Meist handelt es sich um allgemeine Aussagen in FAQs oder Vorstellungen ("About") und eher selten im Rahmen einer Policy. Eine Gemeinsamkeit konnte nicht festgestellt werden. Die angebotenen Sicherungsmaßnahmen sind vielfältig; mehrfach erwähnt werden ortsverteilte Aufbewahrung, redundante Datenhaltung, Bandsicherung, und Format- oder Datenträger Migration. Nur wenige FDR äußern sich zur genauen Sicherungsdauer der anvertrauten Forschungsdaten und nur eines gibt eine explizite Dauer an. Eine echte Nachfolgeregelung gibt keines der Repositorien bekannt. Für weniger als die Hälfte wird wenigstens angegeben, dass die Thematik bekannt ist und eine Übergabe der Daten an unspezifizierte Repositorien geplant ist.

Keines der betrachteten Repositorien ist zertifiziert. Trotz der diversen Zertifizierungsmöglichkeiten und deutlichen Empfehlungen in der Fachliteratur [RS16, S. 116f.].

K4 Nutzererlebnis und Support

Möglichkeiten das Nutzererlebnis zu verbessern sind für jedes Repository unterschiedlich, was nicht überraschend ist, da die Angebote an die Bedürfnisse der Nutzerx angepasst sein müssen.

Die Bedienung der Suchfunktion auf den Webseiten verhält sich mehrheitlich entsprechend den Erwartungen, die durch die regelmäßige Benutzung von Bibliothekskatalogen und Suchmaschinen entstehen. Ebenso werden überwiegend Symbole und Icons genutzt die im Internet üblich sind. Die im Idealfall flache Lernkurve wird in vielen der Repositorien durch die erwähnte fehlende Suchhilfe beeinträchtigt, dies besonders weil eine gezielte Suche nach Daten meist den einzigen Zugang bietet. Keines der betrachteten Repositorien ist barrierefrei gestaltet. Nicht einmal die niedrigste Konformitätsstufe wurde annähernd erreicht.

Im Bereich der Datenvisualisierung wurden wenige Angaben gemacht und keine Gemeinsamkeit gefunden. Dies ist jedoch logisch nachvollziehbar, da die Daten und die Bedürfnisse der Nutzerx je nach wissenschaftlicher Disziplin und Arbeitsweise unterschiedlich sind. Vorschaufunktionen für Bilder werden in der Regel geboten, und in der Hälfte der Repositorien auch für Videos in Form von Online-Streaming im Browser ermöglicht. Die Nutzerx der anderen Repositorien müssen anhand der Metadaten entscheiden ob ein Download der Forschungsdaten sinnvoll erscheint.

Hilfeseiten und Frequently Asked Questions stellen das überwiegende Supportangebot dar. Er-

gänzt werden die verfügbaren Informationen durch Seiten zu einzelnen Themen, in einem Fall liegen für Nutzerx nur Themenseiten vor. Schulungsangebote (Online oder vor Ort) werden in der Hälfte der Repositorien angeboten, dabei ist zu beachten das externe Nutzerx von Schulungen vor Ort nicht profitieren können. Supportanfragen können an alle FDR gesandt werden, die eine Hälfte bietet ein Kontaktformular, die andere eine E-Mail Adresse. Als Nutzer- und Mitarbeiterfreundlicher wird das im IT-Support etablierte Formular eingeschätzt. Die Felder geben Nutzerx einen Anhalt über die Informationen die zur Lösung der meisten Probleme benötigt werden und ersparen Rückfragen. Nur ein Repository bittet seine Nutzerx darum, gefundene Fehler in den FD oder MD zu melden. Dies beinhaltet Tippfehler genauso wie inhaltliche Fehler, die nachnutzenden Personen wahrscheinlich eher auffallen als Datengeberx oder Mitarbeiterx.

Eine Beratung zum oder Hilfestellung im Management von Forschungsdaten erfolgt in einer Mehrheit der analysierten Repositorien. Die Themen und Tiefe der Angebote unterscheiden sich jedoch in allen Fällen. Der Grund dafür wird in den verschiedenen bedienten Communities und deren eigenen Bedürfnissen und Arbeitsweisen vermutet. Auch die Standards und Richtlinien des FDM auf die in den meisten Fällen verwiesen wird überschneiden sich nicht, obwohl dies durchaus disziplinunabhängig möglich wäre.

K5 Organisation und Infrastruktur

Die untersuchten Repositorien werden überwiegend von einer Bibliothek inhaltlich betreut, für die technische Unterstützung wird in wenigen Fällen die IT-Abteilungen des Trägers genannt. Zur Langfristigkeit der Organisationsform oder des Trägers wird nur eine Aussage getroffen. Alle analysierten FDR sind jedoch an etablierten Hochschulen oder Instituten angesiedelt, welches laut re3data bereits eine Aussage zur Stabilität darstellt [Rüc+15, S. 4]. Die Träger werden auf allen untersuchten Webseiten deutlich genannt, es sind hauptsächlich Universitäten. Zum Gewinnstreben des Repositoriums gibt nur ein Repository Auskunft, dieses ist Gemeinnützig. Die Anzahl der Mitarbeiter wird nur auf zwei Webseiten preisgegeben, ermittelbar ist sie in der Regel nur über die zuständigen Abteilungen. Qualifikationen werden von der Hälfte der FDR über die Stellenbezeichnungen angedeutet und in keinem Fall detaillierter beschrieben.

Über die Hälfte der Repositorien entwickelt ihre eigene technische Infrastruktur weiter oder nimmt an Infrastrukturprojekten teil, die sich mit der Entwicklung von Repositoriensoftware u.ä. befassen. An Forschungsprojekten die Forschungsdatenmanagement im Allgemeinen oder verwandte Themen in einzelnen Disziplinen thematisieren beteiligen sich die Hälfte der analysierten FDR. Außerdem arbeiten drei Repositorien an der Weiterentwicklung von FDM- Standards mit. Sechs der Repositorien listen diverse nationale, internationale und fachspezifische Forschungsprojekte auf ihren Webseiten auf.

Alle Webseiten nennen mindestens allgemeine technische Grundlagen für ihre Repositorien, mehrfach werden Fedora und EPrints als Basis genannt. Die Ladezeiten der Webseiten sind in der Regel nicht auffällig schnell oder langsam im Vergleich zu anderen fachlichen oder alltäglichen Webseiten. Aussagen zur Ausfallsicherheit wurden zwei Mal gemacht. In beiden Fällen werden Ausfallquoten durch verteilte Systeme reduziert. Backup-Strategien werden in keinem Fall ausführlich dargestellt. Die drei Repositorien die Angaben dazu machen nennen mehrfach die regelmäßige Sicherung auf Magnetband. Nur ein FDR gibt an eine Überwachung der Hardwarekomponenten durchzuführen. Die technische Dokumentation der Repositorien ist nur auf zwei Webseiten erwähnt. Alle Repositorien die technische Informationen bereitstellen, nennen überwiegend offene Software.

K6 Gebühren und Finanzierung

Die Hälfte der Repositorien geben explizit an, dass das Angebot kostenfrei genutzt werden kann. Für die weiteren war mindestens Zugang zu Daten im Rahmen der Analyse ebenfalls

kostenfrei möglich. Der Ingest konnte wegen fehlender Institutionszugehörigkeit in den meisten Fällen nicht geprüft werden. Ein Gebührenmodell liegt in keinem FDR vor, in einem Fall ist es aber in Entwicklung. In keinem Fall wurde ein Vertrag zwischen den Institutionen oder deren Abteilungen und dem FDR erwähnt, die untersuchten Repositorien und deren Datengeber sind jedoch mehrheitlich in die selbe Trägerorganisation eingebunden wie das FDR.

Keines der analysierten Repositorien verweist öffentlich auf ein langfristiges Konzept zur Finanzierung des Betriebs. Nur in einem Fall konnte eine Budgetschätzung gefunden werden und diese nur in einem Artikel und nicht auf der Webseite des FDR. Keines der FDR bietet eine Prognose für die Finanzierung der Dienste in der Zukunft.

An dieser Stelle sei an das Kapitel 2 erinnert und den Umstand, dass für diverse der untersuchten Kriterien der Umgang mit den Forschungsdaten nicht von deren Art abhängig ist. D.h. textuelle, visuelle, tabellarische und sonstige Forschungsdaten können in vielen Punkten gleich behandelt werden. Die Anfangs identifizierten Punkte die für Videodaten besonders sind werden mit Bezug auf die erfolgte Analyse noch einmal getrennt betrachtet:

K1 Veröffentlichung und Nachnutzung

Datenabgabe: Eine Eingrenzung auf archivfreundliche Formate wird in den betrachteten FDR selten vorgenommen. Empfohlene Formate auf den Webseiten der FDR stimmen eher nicht mit den UK Data Service benannten Formaten [UK oJ] überein. Eine Erwähnung erhält MPEG, welches ohne Version genannt wird.

Metadaten: Keines der FDR legte sein Metadatenschema offen. Die Metadatenschemata DC und METS, die auch für Video empfohlen werden [RB10] und [nes16, S. 52], werden genutzt. Ein anderweitiger Hinweis auf besondere Felder für Videos wurde nicht gefunden, ebensowenig wurde die Einbindung von einem der Metadatenschemata speziell für Videos erwähnt.

Zitierfähigkeit / Verlinkung: Es wurde keine Möglichkeit gefunden einen bestimmten Zeitpunkt in einem Video zu zitieren.

Nutzungsstatistik: Keines der FDR zeigt eigene Nutzungsstatistiken für Anzahl der Aufrufe ("Views") von Videos oder eine Verwendung des COUNTER Code of Practice²⁰¹.

K2 Zugriff und Auffindbarkeit

Recherchemöglichkeiten: Keines der FDR thematisiert Video Retrieval, gleichzeitig wird aber auch das Retrieval für alle anderen Arten von Forschungsdaten auf keiner Webseite erwähnt.

K3 Transparenz und Qualitätssicherung

Sicherstellung langfristiger Zugänglichkeit: Weder zur Problematik schnell wechselnder Dateiformate noch zur verlustreichen Kompression von Video wurden Aussagen getätigt. Migration wird mehrfach genannt, auf die erhöhten Kosten von Formatmigration für Videodaten [Ing17] wird nicht eingegangen.

Usability: In keinem FDR wurden besonders benutzerfreundliche Angebote für Videos festgestellt, auch Annotationsmöglichkeiten für Videos sind nicht vorhanden. Wie bereits dargestellt, wird die Barrierefreiheit für Videos nicht durch Transkripte, Untertitel oder inhaltliche Beschreibungen unterstützt.

Darstellung der Datensätze: Streaming wird in der Hälfte der FDR als Vorschaufunktion für Videos angeboten.

Support zu Repositoriumsaktionen / Beratung zum Forschungsdatenmanagement: In keinem Fall wurden spezielle Informationen zum Umgang mit Videos genannt. Ob diese Inhalt von Präsenz-Schulungen sind konnte durch die Analyse der Webseiten nicht ermittelt werden.

K5 Organisation und Infrastruktur

²⁰¹https://www.projectcounter.org/wp-content/uploads/2017/07/Friendly_Guide_Providers_20170706-1.pdf

Personalausstattung und -qualifizierung: Keines der FDR bietet Informationen zu Mitarbeiterx deren Aufgabenbereich die Behandlung von Videos als Forschungsdaten beinhaltet.

Forschungs- und Entwicklungstätigkeiten: Es wurde für kein FDR eine Teilnahme an Video-relevanten Projekten festgestellt. Viele der FDR zeigen jedoch nur eine Auswahl der Projekte auf ihren Webauftritten.

Technische Infrastruktur: Besondere Maßnahmen für Videos werden nicht erwähnt, ebenso wenig wird die Problematik von Backups für große Datenmengen thematisiert.

K6 Gebühren und Finanzierung

Für Gebühren für Datenabgabe und -nutzung und zur Finanzierung des Betriebs sind keine Besonderheiten von Videodaten bekannt.

Bewertung des Vorgehens

Die Auswahl der Repositorien anhand der re3data Metadaten wird, nicht zuletzt wegen der Fehleranfälligkeit, als problematisch eingeschätzt. Für ähnliche Untersuchungen wird eine Betrachtung von zufällig bestimmten Stichproben empfohlen. Wenn dennoch eine Einschränkung anhand von re3data oder ähnlichen Verzeichnissen vorgenommen werden soll, wird ein Kriterium empfohlen, das Zugang zum Datenupload sicherstellt (in re3data z.B. dataUpload:open), welches die Erhebung der Daten für die Kategorien "Veröffentlichung und Nachnutzung" und "Zugriff und Auffindbarkeit" vereinfacht hätte.

Acht Repositorien werden für den umfangreichen Kriterienkatalog als zu viele empfunden. Eine geringere Anzahl von FDR in höherer Tiefe zu betrachten ist wünschenswert und hätte den Umfang der vorliegenden Arbeit begrenzt. Alternativ kann ein schlankerer Kriterienkatalog, z.B. mit einem Schwerpunkt auf Video-relevanten Aspekten, den Fokus einer solchen Arbeit schärfen. Dies allerdings zu Lasten eines umfassenden Überblickes der durch die Verwendung des gesamten Kataloges gebildet werden konnte.

Die Methode der inhaltlichen Analyse der Webseiten konnte für viele der Kriterien zufriedenstellende Ergebnisse sammeln. Andere Aspekte (darunter Finanzierungs- und Personalthemen) wurden nur sehr verhalten kommuniziert. Hier kann potentiell eine Befragung einer zuständigen Person im Repository vollständiger Aussagen erreichen. Bei Befragungen ist jedoch besonders die externe Abhängigkeit zu beachten, die bei der gewählten Methode sehr viel geringer war.

Durch den Fokus auf Webseiten wird vor allem deutlich welche Informationen explizit gegenüber Nutzerx der Webseite geäußert werden, dies beinhaltet Datengeberx sowie die Daten nachnutzenden Personen.

Hinsichtlich des Fehlers in der Datengrundlage, der die Auswahl der Repositorien auf eine Teilmenge der ursprünglich angepeilten Gesamtzahl reduzierte, wäre eine Untersuchung der hier ausgeschlossenen Repositorien interessant. Vor allem unter Anwendung der Webseitenanalyse und des erweiterten Kriterienkataloges von Steinhoff wäre zu untersuchen ob diese Ergebnisse mit denen dieser Auswahl übereinstimmen und wo Abweichungen bestehen. Ebenso interessant wäre die Untersuchung der bewusst ausgeschlossenen PUSH Repositorien, die ihre Daten aus anderen Quellen aggregieren. Für Repositorien die Daten aggregieren und gleichzeitig Nutzerxuploads zulassen (hier genannt Push/Pull, z.B.: eBird, Internet Archive) wäre vor allem interessant ob eine Abweichung in Kategorien wie "Veröffentlichung und Nachnutzung" oder "Transparenz und Qualitätssicherung" deutlich würde.

Die hier durchgeführte erste Analyse der Daten aus re3data, die unter anderem die Richtigkeit der Angaben aus den Metadaten überprüfte, hätte unter Zuhilfenahme von automatisierten Methoden bzw. Harvesting vorgenommen werden können. Dafür fehlte der Autorin jedoch

die notwendige Infrastruktur (z.B. Server). Ein größer angelegter Abgleich der Aussagen auf re3data und den Webseiten der Repositorien ist ein Thema, das potentiell eine weitere Bachelorarbeit füllen kann.

Die stichprobenartige Untersuchung der unter K2.3 Recherchemöglichkeiten erfragten Suchoperatoren war wegen fehlender Suchhilfen nötig und konnte nur behelfsmäßig durchgeführt werden. Hier sind potentiell weitere Untersuchungen interessant, denn oft waren die Ergebnisse für Tests mit Wildcards und Trunkierung nicht Beweiskräftig. Dies war aber kein zentraler Bestandteil dieser Arbeit und musste deshalb in aller Kürze abgehandelt werden.

Die Unterpunkte des Kriteriums 4.1 Usability (Intuitive Bedienung, Etablierte Darstellungsform und Flache Lernkurve) sind sehr subjektiv und waren mit der Methode nicht befriedigend zu beantworten. Eine Befragung von Nutzerx (qualitativ oder quantitativ) wäre hier angebrachter, aber im Rahmen dieser Arbeit nicht vorgesehen.

Steinhoff erstellte gemeinsam mit den hier nachgenutzten Kriterien auch eine Gewichtung der einzelnen Punkte für verschiedene Stakeholder des FDM, darunter auch die Gruppe Benutzerx deren Perspektive hier eingenommen wurde. Leider konnte die Gewichtung aus Kapazitätsgründen nicht mit den in dieser Arbeit gesammelten Ergebnissen abgeglichen werden um die für Nutzerx besonders zentralen Aspekte höher zu gewichten. Ein Abgleich mit den Gewichtungen der hier gewonnenen Ergebnissen und einer Befragung von Nutzerx einem oder mehrerer Repositorien könnte z.B. eine Priorisierung von diversen Verbesserungsmöglichkeiten erlauben.

8 Fazit

In der Betrachtung multidisziplinärer Repositorien können alle Arten von Daten in der Forschung benutzt werden oder entstehen. Videos sind also nur eine der vielen Formen die Forschungsdaten annehmen können und müssen sich in das große Ganze einfügen. Für Forscher die Videos produzieren oder Personen, die wissenschaftliche Videos nachnutzen möchten, ist es dennoch von Bedeutung, dass die Besonderheiten dieser Ressource in den Forschungsdatenrepositorien beachtet werden.

Wie in Kapitel 7 deutlich wurde, behandeln die analysierten Forschungsdatenrepositorien alle Arten von Forschungsdaten gleich. Von den identifizierten Besonderheiten für die Behandlung von Videos wurde lediglich Streaming als Vorschaufunktion umgesetzt, und dies nur in der Hälfte der Repositorien. Keine der anderen Möglichkeiten (u.a. archivfreundliche Formate, formatspezifisches Metadatenschema, Accessibility durch Untertitel) wurde beachtet.

Gemeinsamkeiten und Unterschiede konnten für audiovisuelle Forschungsdaten nur insofern festgestellt werden, dass sie keine Sonderbehandlung erfahren. Für den Umgang mit allgemeinen Forschungsdaten konnten in den FDR jedoch eine Anzahl von Gemeinsamkeiten und Unterschieden identifiziert werden, sowie diverse Tendenzen, die im vorigen Kapitel zusammengefasst wurden.

Insgesamt wird für die betrachteten FDR deutlich, dass Gebiete auf denen diverse Standards, Guidelines und Best Practice Beispiele zur Verfügung stehen (Langzeitarchivierung, Usability, Finanzierung, Erhebung von Statistiken, Barrierefreiheit, Zertifizierung, typenspezifische Metadaten) besonders große Heterogenität aufweisen. Diese Unterschiede sind nicht durch einen Fokus auf die Community und deren disziplinspezifische Bedürfnisse erklärbar, da sie nicht primär von den Erwartungen und Arbeitsweisen der Disziplin abhängen. Andere auffällige Unterschiede (Unterstützung bei der Eingabe von Metadaten, Spezifika des Forschungsdatenmanagements, Schulungsangebote, Datenvisualisierung, Recherchemöglichkeiten) sind dagegen deutlich disziplinabhängig und müssen für jede Zielgruppe konzipiert oder zumindest angepasst werden. Es wurde jedoch auch deutlich, dass die internationalen Standardisierungsbemühungen für einige Aspekte mindestens Tendenzen zur Akzeptanz zeigen. Hier sind vor allem die Verwendung von Creative Commons Lizenzen, des DOI Systems zur persistenten Identifizierung, von Programmierschnittstellen und des Metadatenschemas Dublin Core zu erwähnen.

Die Behandlung von Forschungsdaten im Allgemeinen (inkl. Videos) ist transparent dargestellt. Viele der Aspekte des umfangreichen Kriterienkataloges konnten zufriedenstellend aus den Webseiten extrahiert werden. Die Informationen sind jedoch auf viele Unterseiten verteilt und in einigen Fällen nur impliziert.

Aspekte zu denen für alle FDR eine Aussage gemacht werden konnte, beinhalten die Verknüpfung von Projekten und Forschungsdaten (K1.6 Verweise auf Veröffentlichungen), den Nutzerkreis der zum Upload berechtigt ist (K2.1 Zugangsformen), das Vorhandensein von Richtlinien (K3.1 Policies), die Übernahme der Verantwortung für die Langzeitarchivierung durch das Repository (K3.5 Sicherstellung langfristiger Zugänglichkeit), die Art der Einbindung in Organisation und Bekanntgabe des Trägers (K5.1 Organisationsform) und Allgemeine Informationen zu Soft- und Hardware, Lade- & Reaktionszeiten und der Verwendung offener Software (K5.4 Technische Infrastruktur).

Weitere vierzehn Aspekte konnten für sieben von acht FDR ermittelt werden, darunter die zugelassenen Formate, Verwendung von persistenten Identifiern, APIs, Maßnahmen der LZA, Vorschaufunktionen usw.

Nur eine Aussage konnte für sieben der Aspekte extrahiert werden, diese vor allem in den Kategorien 5 "Organisation und Infrastruktur" und 6 "Gebühren und Finanzierung", aber auch zum Grund der Löschung eines Datensatzes (K3.4 Löschung) und der Möglichkeit Fehler zu melden (K4.3 Support zu Repositoriumsaktionen).

Zu einigen Aspekten konnten gar keine Informationen gefunden werden. Darunter die Verlinkung von Forschungsdaten und nachnutzenden Publikationen (K1.6 Verweise auf Veröffentlichungen), das Vorhandensein einer standardisierten Erhebung von Statistiken (K1.7 Nutzungsstatistik), eine Anzeige des Grundes für ein Embargo (K2.2 Embargofunktion), die Einbindung von Empfehlungssystemen zur Erhöhung der Serendipität (K2.3 Recherchemöglichkeiten), die Langfristigkeit der Organisationsform und Hinweise zur Gewinnerorientierung (K5.1 Organisationsform), Informationen zum Gebührenmodell oder ob ein Vertragsverhältnis zwischen FDR und Träger besteht (K6.1 Gebühren für die Datenabgabe und -nutzung), sowie die Aspekte Finanzierungskonzept und Finanzierungsprognose (K6.2 Finanzierung des Betriebs).

Die Antwort auf die Frage, ob Nutzerx anhand der Informationen auf den Webseiten entscheiden können, ob sie ihre Videos dort ablegen, ist für sechs der hier untersuchten FDR bereits von Hause aus eingeschränkt, da es sich um institutionelle Repositorien handelt und keine andere Alternative existiert. Auch ob Forschungsdaten abgegeben werden müssen oder nicht, ist meist intern geregelt (z.B. durch eine Forschungsdatenpolicy). Dennoch ist der Vertrauensgewinn gegenüber Nutzerx durch einige Punkte eher eingeschränkt. Darunter fallen z.B. fehlende Zertifizierung, nur vereinzelte Aussagen zum langfristigen Bestehen der FDR, wenige und unstrukturierte Aussagen zur Sicherstellung der langfristigen Zugänglichkeit der Forschungsdaten und der vorhandenen technischen Infrastruktur.

8.1. Ausblick

Um die Transparenz zu erhöhen und den Nutzerx die Suche nach Informationen zum FDR zu erleichtern, sollten Hilfeseiten und Richtlinien explizit als solche benannt werden. Wenn eine thematische Unterteilung auf mehrere Seiten erfolgt, sollten die Themen von einer gemeinsamen Stelle erreichbar sein um einen Überblick zu ermöglichen. Der Übergang von der Suchoberfläche zu den Hilfeseiten sollte mit einem Klick möglich sein.

Das Löschen von Datensätzen muss weitaus deutlicher geregelt werden. Die analysierten Vorgehensweisen können trotz persistenten Identifiern zu Broken Links führen. Wenn außerdem nicht über den Grund einer Löschung informiert wird, können Nutzerx z.B. nicht bewerten ob ein Datensatz wegen wissenschaftlichen Fehlverhaltens o. ä. gelöscht wurde.

Die Verantwortung für die Richtigkeit der Daten muss flächendeckend und deutlich kommuniziert werden. Im Problemfall darf kein Zweifel entstehen, ob FDR oder Datengeberx rechtlich zu belangen sind.

Die technische Grundlage muss soweit kommuniziert werden, dass Nutzerx daran festmachen können ob es sich um aktuelle Soft- und Hardware handelt, die regelmäßig überprüft, gewartet und erneuert wird. Der Umfang und die Granularität der Details sollten dem vermuteten technischen Verständnis der Nutzerx angepasst werden.

Die Anzeige von Publikationen, die die Forschungsdaten nachnutzen, sollte auf Datensatzeben implementiert werden. Technisch möglich ist dies bereits durch die persistente Verlinkung von Forschungsdaten mit DOIs. Eine solche Anzeige würde Forscherx und Nachnutzerx enger verknüpfen und eventuell weitere Nachnutzerx dazu animieren, die Forschungsdaten für eigene Untersuchungen in Betracht zu ziehen. Eine manuelle Verlinkung durch Mitarbeiterx des FDR ist auch möglich, aber weder flächendeckend noch ökonomisch sinnvoll.

Die Metadaten schemata der untersuchten Repositorien sind nicht öffentlich, oft werden aber Metadatenstandards genannt die ganz oder teilweise nachgenutzt wurden. Die Betreiberx der FDR gehen davon aus, dass Forscherx und Nachnutzerx wissen, welche Standards es gibt oder ein Interesse haben sich dazu selbstständig zu informieren. Eine Einbindung von typenspezifischen Metadatenstandards z.B. zur Beschreibung von Videos kann Nutzerx dabei unterstützen ihre Forschungsdaten vollständiger zu beschreiben.

Die Metadaten sollten auf den Webseiten deutlich lizenziert werden. Im Idealfall bindungslos um eine Nutzung der Metadaten in anderen Kontexten schwellenlos zu ermöglichen. Eine aktive Lieferung von Metadaten an Zitationsdatenbanken erhöht die Sichtbarkeit der Forschungsdaten und damit potentiell den Impact für Datengeberx.

Vor allem bei der Untersuchung der Barrierefreiheit wurde deutlich, dass für Videos eine besondere Behandlung nötig ist um die Nachnutzung für alle Nutzerx zu ermöglichen. Durch diese Schwellen werden Nutzerx aller acht Repositorien mit jeder Art von körperlicher oder geistiger Beeinträchtigung de facto von der Nachnutzung der Forschungsdaten ausgeschlossen. Wenn außerdem bedacht wird, dass der Zugang zu Daten eine der Hauptaufgaben von Repositorien ist (erinnert sei an die Definition für FDR aus [Kin+17, S. 2]), ist eine barrierefreie Gestaltung der Webseiten unabdingbar. Dies gilt exponentiell für die Gestaltung der Suchoberfläche.

Um Forscherx, die Videos und andere audiovisuelle Forschungsdaten in einem Repository ablegen wollen, zu unterstützen und die Forschungsdaten so findbar und nachnutzbar wie möglich zu machen, stehen diverse Standards und Guidelines zur Verfügung. Forschungsdatenrepositorien, die die Abgabe und Nachnutzung von Videos fördern möchten, sollten sich an den vorliegenden Ressourcen bedienen.

9 Verzeichnisse

9.1. Abbildungsverzeichnis

2.1. Auswahl der Repositorien	6
2.2. Anzahl von FDR nach <code>dataAccessType</code> und <code>dataAccessRestrictions</code> in <code>re3data</code>	7

9.2. Tabellenverzeichnis

2.1. Ausschlussgründe und deren Anzahl in der ersten Analyse der Repositorien	8
A.1. Liste der Repositorien aus <code>re3data</code>	viii
A.2. Ergebnisse der ersten Analyse: Anzahl an AV	ix
A.3. Ergebnisse der ersten Analyse: Prozentualer Anteil an Videos im Vergleich zu allen Ergebnissen	ix
A.4. Ergebnisse der ersten Analyse: Push/Pull Prüfung	x
A.5. Vergleich K1 Veröffentlichung und Nachnutzung	lxv
A.6. Vergleich K2 Zugriff und Auffindbarkeit	lxvi
A.7. Vergleich K3 Transparenz und Qualitätssicherung	lxvii
A.8. Vergleich K4 Nutzererlebnis und Support	lxvii
A.9. Vergleich K5 Organisation und Infrastruktur	lxviii
A.10. Vergleich K6 Gebühren und Finanzierung	lxix

Literatur

- [Abt Ja] Abteilung Phaidra-UNIDAM. *Anbindung an Projekte - nationale und interne Projekte*. Wien, o. J. URL: <http://datamanagement.univie.ac.at/ueber-phaidra/anbindung-an-projekte/> (besucht am 26. 12. 2017).
- [Abt Jb] Abteilung Phaidra-UNIDAM. *Forschungsdatenmanagement (Research Data Management)*. Wien, o. J. URL: <https://datamanagement.univie.ac.at/rdm/> (besucht am 26. 12. 2017).
- [Abt Jc] Abteilung Phaidra-UNIDAM. *Metadaten - mehr als Daten über Daten*. Wien, o. J. URL: <https://datamanagement.univie.ac.at/rdm/metadaten/> (besucht am 26. 12. 2017).
- [Abt Jd] Abteilung Phaidra-UNIDAM. *Netzwerk nachhaltige Softwareentwicklung*. Wien, o. J. URL: <https://datamanagement.univie.ac.at/rdm/netzwerk-nachhaltige-softwareentwicklung/> (besucht am 26. 12. 2017).
- [Abt Je] Abteilung Phaidra-UNIDAM. *Phaidra Systeme*. Wien, o. J. URL: <http://datamanagement.univie.ac.at/ueber-phaidra/phaidra-systeme/> (besucht am 26. 12. 2017).
- [Abt Jf] Abteilung Phaidra-UNIDAM. *Policy von Phaidra*. Wien, o. J. URL: <http://datamanagement.univie.ac.at/ueber-phaidra/policy-von-phaidra/> (besucht am 26. 12. 2017).
- [Abt Jg] Abteilung Phaidra-UNIDAM. *Über Phaidra*. Wien, o. J. URL: <http://datamanagement.univie.ac.at/ueber-phaidra/> (besucht am 26. 12. 2017).
- [Abt Jh] Abteilung Phaidra-UNIDAM. *Unterstützte Formate*. Wien, o. J. URL: <https://backend.univie.ac.at/index.php?id=23004&L=0> (besucht am 26. 12. 2017).
- [Abt Ji] Abteilung Phaidra-UNIDAM. *Zugang zu Phaidra*. Wien, o. J. URL: <http://datamanagement.univie.ac.at/ueber-phaidra/zugang-zu-phaidra/> (besucht am 26. 12. 2017).
- [Abt08] Abteilung Phaidra-UNIDAM. *Phaidra Tutorial 4 - Metadateneditor*. Wien, 2008. URL: https://phaidra.univie.ac.at//static/tutorials/tutorial_4_metadaten.pdf (besucht am 26. 12. 2017).
- [Ada+17] S. Adams Becker u. a. *NMC Horizon Report: 2017 Library Edition*. Austin, Texas: The New Media Consortium, 2017. ISBN: 978-0-9986242-8-0.
- [AM14] Lucy Ayre und Natalia Madjarevic. *20 ways to reuse repository content [Poster]*. Open Repositories 2014, 9-13 June 2014. Helsinki, 2014. URL: <http://eprints.lse.ac.uk/id/eprint/57104> (besucht am 30. 12. 2017).
- [Ami J] Seyavash Amini. *User Konzept und Nutzungsbedingungen der Dienste Phaidra und Tempus*. o. J. URL: https://phaidra.univie.ac.at/terms_of_use/show_terms_of_use (besucht am 27. 12. 2017).
- [AN11] Andreas Aschenbrenner und Heike Neuroth. "Forschungsdaten-Repositorien". In: *Handbuch Forschungsdatenmanagement*. Hrsg. von Stephan Büttner, Hans-Christoph Hobohm und Lars Müller. Bad Honnef: Bock + Herchen, 2011, S. 101–114. ISBN: 978-3-88347-283-6. URL: <urn:nbn:de:kobv:525-opus-2328>.

- [And13] R. Anderson. *The Moab Design for Digital Object Versioning*. 2013. URL: <http://purl.stanford.edu/vt105qd7230> (besucht am 14. 12. 2017).
- [Aso17] Lavanya Asokan. *Launch of NTU Open Access Research Data Repository DR-NTU (Data)*. Singapur, 2017. URL: <https://blogs.ntu.edu.sg/lib-scholarlycomm/?p=19143> (besucht am 21. 12. 2017).
- [Ass+16] Massimiliano Assante u. a. "Are Scientific Data Repositories Coping with Research Data Publishing?" In: *Data Science Journal* 15 (2016). ISSN: 1683-1470. DOI: 10.5334/dsj-2016-006. URL: <http://datascience.codata.org/articles/10.5334/dsj-2016-006/galley/593/download/>.
- [Ayr13] Lucy Ayre. "Libraries can embrace the use of altmetrics as a means to strengthen the functionality of institutional repositories". In: *Impact of Social Sciences Blog* (2013). URL: <http://eprints.lse.ac.uk/id/eprint/51834> (besucht am 30. 12. 2017).
- [Bad16] Tabea Bader. "Open-Access-Repositorien in der Schweiz und Österreich - Auswertung des 2014 Census on Open Access Repositories". In: *LIBREAS. Library Ideas* 29 (2016), S. 113–131. URL: <http://libreas.eu/ausgabe29/13bader/>.
- [Bad17] Tabea Bader. "Open Access Repositorien in Skandinavien - Eine Analyse der Situation in Schweden, Norwegen, Dänemark und Finnland". Magisterarb. Berlin: Humboldt-Universität zu Berlin, 2017. DOI: 10.18452/14296. URL: <https://edoc.hu-berlin.de/handle/18452/14948>.
- [Bau+15] Bruno Bauer u. a. *Forschende und ihre Daten. Ergebnisse einer österreichweiten Befragung – Report 2015*. 2015. DOI: 10.5281/zenodo.32043. URL: <https://doi.org/10.5281/zenodo.32043> (besucht am 27. 12. 2017).
- [Bec14] Pascal-Nicolas Becker. "Repositorien und das Semantic Web - Repositorieninhalte als Linked Data bereitstellen". Magisterarb. Berlin: Freie Universität Berlin, 2014. DOI: 10.14279/depositonce-5015.
- [BF J] S. Blumesberger und A. Ferus. *Netzwerk für RepositorienmanagerInnen*. Wien, o. J. URL: <https://datamanagement.univie.ac.at/rdm/netzwerk-fuer-repositorienmanagerinnen/> (besucht am 27. 12. 2017).
- [BG17] S. Blumesberger und R. Ganguly. *Data Management und Data Archiving an der Universität Wien [Folien]*. 2017. URL: http://datamanagement.univie.ac.at/fileadmin/user_upload/p_phaidraservice/Schulungsunterlagen_Datenarchivierung_RG_SB.pdf (besucht am 27. 12. 2017).
- [BHM11] Stephan Büttner, Hans-Christoph Hobohm und Lars Müller, Hrsg. *Handbuch Forschungsdatenmanagement*. Bad Honnef: Bock + Herchen, 2011. ISBN: 978-3-88347-283-6. URL: <http://nbn-resolving.de/urn:nbn:de:kobv:525-opus-2412>.
- [Blu15] Susanne Blumesberger. "Die Welt der Metadaten im Universum von Repositorien". In: *Mitteilungen der VÖB* 68.3/4 (2015), S. 515–528. ISSN: 1022-2588. URL: <https://ojs.univie.ac.at/index.php/voebm/article/view/1295>.
- [Blu16] S. Blumesberger. *Wie kann Phaidra Ihre Forschung unterstützen?* Wien, 2016. URL: http://datamanagement.univie.ac.at/fileadmin/user_upload/p_phaidraservice/Foschung_unterstuetzen.pdf (besucht am 27. 12. 2017).

- [BPZ16] Susanne Blumesberger, José Luis Preza und Alexander Zartl. *Umgang mit Metadaten in Repositorien - Eine österreichweite Umfrage*. Techn. Ber. e-infrastructures austria, 2016. URL: <https://hdl.handle.net/11353/10.441216>.
- [Bro Ja] Brown Computing & Information Services. *Closed Captioning for Digital Content*. Providence, RI, o. J. URL: <https://it.brown.edu/services/type/closed-captioning-digital-content> (besucht am 21. 12. 2017).
- [Bro Jb] Brown Digital Repository. *About the Brown Digital Repository*. Providence, RI, o. J. URL: <https://repository.library.brown.edu/studio/about/> (besucht am 21. 12. 2017).
- [Bro Jc] Brown Digital Repository. *Brown Digital Repository APIs*. Providence, RI, o. J. URL: <https://repository.library.brown.edu/studio/api-docs/> (besucht am 21. 12. 2017).
- [Bro Jd] Brown Digital Repository. *Digital Object Identifiers*. Providence, RI, o. J. URL: <https://repository.library.brown.edu/studio/doi/> (besucht am 21. 12. 2017).
- [Bro Je] Brown Digital Repository. *Frequently Asked Questions [BDR]*. Providence, RI, o. J. URL: <https://repository.library.brown.edu/studio/faq/> (besucht am 21. 12. 2017).
- [Bro Jf] Brown Digital Repository. *Policies*. Providence, RI, o. J. URL: <https://repository.library.brown.edu/studio/policies/> (besucht am 21. 12. 2017).
- [Bro Jg] Brown Library - Center for Digital Scholarship. *Activities*. Providence, RI, o. J. URL: <https://library.brown.edu/cds/activities/#bdr> (besucht am 21. 12. 2017).
- [Bro Jh] Brown Library - Center for Digital Scholarship. *Research*. Providence, RI, o. J. URL: <https://library.brown.edu/cds/research/> (besucht am 21. 12. 2017).
- [Bro Ji] Brown Library - Digital Production Services. *Digital Production Services*. Providence, RI, o. J. URL: <https://library.brown.edu/dps/> (besucht am 21. 12. 2017).
- [Bro Jj] Brown Library - Digital Technology. *Digital Technologies Strategic Goals, 2017-18*. Providence, RI, o. J. URL: https://library.brown.edu/dt/DT_Goals_2017-18.pdf (besucht am 21. 12. 2017).
- [Bro Jk] Brown University Library. *Deposit Research Data into a Repository*. Providence, RI, o. J. URL: https://library.brown.edu/info/data_management_depositing (besucht am 21. 12. 2017).
- [Bro Jl] Brown University Library. *Digital Technologies*. Providence, RI, o. J. URL: <https://library.brown.edu/dt/> (besucht am 31. 12. 2017).
- [BZ17] Susanne Blumesberger und Alexander Zartl. "Umgang mit Metadaten in Repositorien - Eine österreichweite Umfrage". In: *Mitteilungen der VÖB* 70.2 (2017), S. 249–273. URL: <http://hdl.handle.net/10760/31931>.
- [Cai16a] Ben Cail. *IIIF and the BDR*. Providence, RI, 2016. URL: <https://library.brown.edu/DigitalTechnologies/iiif-and-the-bdr/> (besucht am 21. 12. 2017).
- [Cai16b] Ben Cail. *Loris Image Server Deployment*. Providence, RI, 2016. URL: <https://library.brown.edu/DigitalTechnologies/loris-image-server/> (besucht am 21. 12. 2017).

- [Cai17a] Ben Cail. *Fedora Functionality*. Providence, RI, 2017. URL: <https://library.brown.edu/DigitalTechnologies/fedora-functionality/> (besucht am 21. 12. 2017).
- [Cai17b] Ben Cail. *Storing Embargo Data in Fedora*. Providence, RI, 2017. URL: <https://library.brown.edu/DigitalTechnologies/storing-embargo-data-in-fedora/> (besucht am 21. 12. 2017).
- [Cai17c] Ben Cail. *Upgrades and Architecture changes in the BDR*. Providence, RI, 2017. URL: <https://library.brown.edu/DigitalTechnologies/upgrades-and-architecture-changes-in-the-bdr/> (besucht am 21. 12. 2017).
- [Cal+09] Ben Caldwell u. a., Hrsg. *Richtlinien für barrierefreie Webinhalte (WCAG) 2.0: Autorisierte deutsche Übersetzung der Web Content Accessibility Guidelines (WCAG) 2.0*. 2009. URL: <https://www.w3.org/Translations/WCAG20-de/#guidelines> (besucht am 14. 12. 2017).
- [Cra11] Tom Cramer. *The Stanford Digital Repository: Foliensatz*. 2011. URL: http://web.stanford.edu/group/dlss/pasig/PASIG_May2011/IntroductionRecap/Cramer_SDR_for_PASIG_RWC_May_2011.pdf (besucht am 22. 12. 2017).
- [DAN10] DANS. *Data seal of approval: Quality guidelines for digital research data in the Netherlands*. Den Haag: DANS, 2010. ISBN: 978-9490-531-027.
- [Deg14] Kurt Deggeller. *Bestandserhaltung audiovisueller Dokumente*. Praxiswissen. Berlin, Boston: DeGruyter, 2014. ISBN: 978-3-11-028944-2.
- [Egy+17] Attila L. Egyedi u. a. "Embracing Semantic Technology for Better Metadata Authoring in Biomedicine". In: *Proceedings of the 10th International Conference on Semantic Web Applications and Tools for Health Care and Life Sciences (SWAT4LS 2017)*. SWAT4LS 10. Rome, 2017, S. 1202–1216. URL: http://www.swat4ls.org/wp-content/uploads/2017/11/SWAT4LS-2017_paper_14.pdf.
- [Eid12] Horst Eidenberger. *Handbook of Multimedia Information Retrieval*. Wien: atpress, 2012. ISBN: 978-3-848-22283-4.
- [Fla17] Dimity Flanagan. *LSE's "request a copy" service: widening access to research both within and beyond academia*. 2017. URL: <http://blogs.lse.ac.uk/impactofsocialsciences/2017/10/25/lse-request-a-copy-service-widening-access-to-research-both-within-and-beyond-academia/> (besucht am 30. 12. 2017).
- [Fra17] Cordula Franzke. "Repositorien für Forschungsdaten am Beispiel der Digital Humanities im nationalen und internationalen Vergleich - Potentiale und Grenzen". In: *Perspektive Bibliothek* 6.1 (2017), S. 2–33. DOI: [10.11588/pb.2017.1.42164](https://doi.org/10.11588/pb.2017.1.42164).
- [Fro13] Hannah Frost. *Depositing to the Stanford Digital Repository [Video]*. Providence, RI, 2013. URL: <https://www.youtube.com/watch?v=F1T7guVkG8Q> (besucht am 14. 12. 2017).
- [Fro15] Hannah Frost. *What you need to know about the redesigned SDR PURL page*. Providence, RI, 2015. URL: <https://library.stanford.edu/blogs/digital-library-blog/2015/11/what-you-need-know-about-redesigned-sdr-purl-page> (besucht am 21. 12. 2017).

- [Gan17] R. Ganguly. "Phaidra". In: *Digitale Archivierung und Bereitstellung von AV-Medien*. Books on Demand, 2017. Kap. 4.7, S. 109–112. ISBN: 9783744851961.
- [GEO Ja] GEOMAR Helmholtz-Zentrum für Ozeanforschung. *FAQs [OceanRep]*. Kiel, o. J. URL: <http://oceanrep.geomar.de/faqs.html> (besucht am 21. 12. 2017).
- [GEO Jb] GEOMAR Helmholtz-Zentrum für Ozeanforschung. *OceanRep Hilfe für die Suche*. Kiel, o. J. URL: <http://oceanrep.geomar.de/help/> (besucht am 21. 12. 2017).
- [GEO Jc] GEOMAR Helmholtz-Zentrum für Ozeanforschung. *Repository Policies*. Kiel, o. J. URL: <http://oceanrep.geomar.de/policies.html> (besucht am 21. 12. 2017).
- [GEO Jd] GEOMAR Helmholtz-Zentrum für Ozeanforschung. *Statements*. Kiel, o. J. URL: <http://oceanrep.geomar.de/statements.html> (besucht am 21. 12. 2017).
- [GEO15] GEOMAR Helmholtz-Zentrum für Ozeanforschung. *Forschungsdaten*. Kiel, 2015. URL: <https://www.geomar.de/zentrum/einrichtungen/bibliothek/service-fuer-autoren/forschungsdaten/#c21354> (besucht am 21. 12. 2017).
- [GEO17] GEOMAR Helmholtz-Zentrum für Ozeanforschung. *GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel*. Kiel, 2017. URL: <https://www.geomar.de/zentrum/> (besucht am 21. 12. 2017).
- [GEOoJ] GEOMAR Helmholtz-Zentrum für Ozeanforschung. *Willkommen im OceanRep - GEOMAR Repository*. Kiel, o.J. URL: <http://oceanrep.geomar.de/> (besucht am 21. 12. 2017).
- [GHK10] David Giarretta, Hank Harmsen und Christian Keitel. *Memorandum of Understanding: to create a European Framework for Audit and Certification of Digital Repositories*. 2010. URL: <http://www.trusteddigitalrepository.eu/Memorandum%20of%20Understanding.html>.
- [GMH16] Nancy-Diana Gómez, Eva Méndez und Tony Hernández-Pérez. "Social Sciences and Humanities Research Data and Metadata: A Perspective From Thematic Data Repositories". In: *El profesional de la información* 25.4 (2016), S. 545–555. ISSN: 1699-2407. URL: <http://hdl.handle.net/10760/30054>.
- [GMR09] Ann Green, Stuart Macdonald und Robin Rice. *Policy-making for Research Data in Repositories: a Guide*. 2009. URL: <https://www.coar-repositories.org/files/guide.pdf>.
- [Hod14] Amy E. Hodge. *Happy Anniversary SDR!* Providence, RI, 2014. URL: <https://library.stanford.edu/blogs/digital-library-blog/2014/12/happy-anniversary-sdr> (besucht am 22. 12. 2017).
- [Hod17] Amy E. Hodge. *A better home for your scholarly work: the Stanford Digital Repository*. Providence, RI, 2017. URL: <https://library.stanford.edu/blogs/digital-library-blog/2017/04/better-home-your-scholarly-work-stanford-digital-repository> (besucht am 21. 12. 2017).
- [HTP14] Chung-Yi Hou, Cheryl A. Thompson und Carole L. Palmer. "Profiling open digital repositories in the atmospheric and climate sciences: An initial survey". In: *Proceedings of the American Society for Information Science and Technology* 51.1 (2014), S. 1–4. ISSN: 1550-8390. DOI: [10.1002/meet.2014.14505101121](https://doi.org/10.1002/meet.2014.14505101121). URL: <http://dx.doi.org/10.1002/meet.2014.14505101121>.

- [Ing17] Caroline Ingram. *Audiovisual research data*. 2017. URL: <https://www.jisc.ac.uk/guides/audiovisual-research-data> (besucht am 12. 01. 2018).
- [loa17] Alexander Ioannidis. *Version field launched!* Genf, 2017. URL: <http://blog.zenodo.org/2017/11/02/version-field/> (besucht am 29. 12. 2017).
- [Joh+17] Lisa R. Johnston u. a. "Data Curation Network: How Do We Compare? A Snapshot of Six Academic Library Institutions' Data Repository and Curation Services". In: *Journal of eScience Librarianship* 6.1 (2017), e1102. DOI: [10.7191/jeslib.2017.1102](https://doi.org/10.7191/jeslib.2017.1102).
- [Kin+17] Maxi Kindling u. a. "The Landscape of Research Data Repositories in 2015: A re3data Analysis". In: *D-Lib Magazine* 23.3/4 (2017). DOI: [10.1045/march2017-kindling](https://doi.org/10.1045/march2017-kindling).
- [KKM13] Daniela Koudela, Klaus Köhler und Ralph Müller-Pfefferkorn. *Projekt RADIESCHEN - Rahmenbedingungen einer disziplinübergreifenden Forschungsdateninfrastruktur - Report Technik*. 2013. DOI: [10.2312/RADIESCHEN_004](https://doi.org/10.2312/RADIESCHEN_004). URL: <http://gfzpublic.gfz-potsdam.de/pubman/item/escidoc:117050> (besucht am 03. 01. 2018).
- [Lam J] Lamont-Doherty Earth Observatory. *MGDS: Marine Geoscience Data System*. New York, NY, o. J. URL: <http://www.ldeo.columbia.edu/research/marine-geology-geophysics/mgds-marine-geoscience-data-system> (besucht am 21. 12. 2017).
- [LF16] Lucy Lambe und Dimity Flanagan. *International Open Access Week 2016: your library can help open up your research to the world*. LSE Impact Blog. 2016. URL: <http://blogs.lse.ac.uk/impactofsocialsciences/2016/10/27/international-open-access-week-2016-your-library-can-help-open-up-your-research-to-the-world/> (besucht am 02. 01. 2018).
- [Lon J] London School of Economics and Political Science. *Terms of use [LSE]*. London, o. J. URL: <http://www.lse.ac.uk/lse-information/terms-of-use> (besucht am 21. 12. 2017).
- [LSE Ja] LSE Library Services. *Digital Object Identifier (DOI) service*. London, o. J. URL: <http://www.lse.ac.uk/library/research-support/digital-object-identifier-doi-service> (besucht am 21. 12. 2017).
- [LSE Jb] LSE Library Services. *File formats*. London, o. J. URL: <http://www.lse.ac.uk/library/research-support/research-data-management/file-formats> (besucht am 30. 12. 2017).
- [LSE Jc] LSE Library Services. *ORCID at LSE: Connecting research and researchers*. London, o. J. URL: <http://www.lse.ac.uk/library/research-support/orcid> (besucht am 21. 12. 2017).
- [LSE Jd] LSE Library Services. *Research Data Management*. London, o. J. URL: <http://www.lse.ac.uk/library/research-support/research-data-management> (besucht am 30. 12. 2017).
- [LSE Je] LSE Research Online. *LSE Research Online: [FAQ]*. London, o. J. URL: <http://eprints.lse.ac.uk/faq.html> (besucht am 21. 12. 2017).
- [LSE Jf] LSE Research Online. *LSE Research Online Takedown Policy*. London, o. J. URL: <http://eprints.lse.ac.uk/takedown.html> (besucht am 21. 12. 2017).

- [LSE Jg] LSE Research Online. *OAI 2.0 Request Results*. London, o. J. URL: <http://eprints.lse.ac.uk/cgi/oai2?verb=Identify> (besucht am 21. 12. 2017).
- [LSE14] LSE Institute of Public Affairs. *LSE Research Online*. London, 2014. URL: <http://eprints.lse.ac.uk/id/eprint/80076> (besucht am 30. 12. 2017).
- [Mar Ja] Marine Geoscience Data System. *Acknowledgments*. New York, NY, o. J. URL: <http://www.marine-geo.org/about/legal.php#copyright> (besucht am 21. 12. 2017).
- [Mar Jb] Marine Geoscience Data System. *Data Submission Guidelines*. New York, NY, o. J. URL: <http://www.marine-geo.org/submit/guidelines.php> (besucht am 21. 12. 2017).
- [Mar Jc] Marine Geoscience Data System. *File Uploader - Tips and Help*. New York, NY, o. J. URL: http://www.marine-geo.org/help/contribute_help.php (besucht am 21. 12. 2017).
- [Mar Jd] Marine Geoscience Data System. *How to Contribute Data*. New York, NY, o. J. URL: <http://www.marine-geo.org/submit/> (besucht am 21. 12. 2017).
- [Mar Je] Marine Geoscience Data System. *Marine-Geo Digital Library*. New York, NY, o. J. URL: <http://www.marine-geo.org/library/> (besucht am 21. 12. 2017).
- [Mar Jf] Marine Geoscience Data System. *MGDS Data Model Description*. New York, NY, o. J. URL: <http://www.marine-geo.org/about/datamodel.php> (besucht am 21. 12. 2017).
- [Mar Jg] Marine Geoscience Data System. *MGDS Overview*. New York, NY, o. J. URL: <http://www.marine-geo.org/about/overview.php> (besucht am 21. 12. 2017).
- [Mar Jh] Marine Geoscience Data System. *MGDS Web Services*. New York, NY, o. J. URL: http://www.marine-geo.org/tools/web_services.php (besucht am 21. 12. 2017).
- [Mar Ji] Marine Geoscience Data System. *Search Help*. New York, NY, o. J. URL: http://www.marine-geo.org/help/search_help.php (besucht am 21. 12. 2017).
- [Mar Jj] Marine Geoscience Data System. *Terms of Use [MGDS]*. New York, NY, o. J. URL: http://www.marine-geo.org/about/terms_of_use.php (besucht am 21. 12. 2017).
- [Mül15] Dennis Müller. "Art und Verteilung von PDF-Formaten auf wissenschaftlichen deutschen Repositorien mit Hinblick auf Langzeitarchivierung". Magisterarb. Stuttgart: Hochschule der Medien, 2015. URL: <https://ub-madoc.bib.uni-mannheim.de/38775>.
- [Nan15] Nanyang Technological University. *NTU Research Data Policy*. Singapur, 2015. URL: <http://research.ntu.edu.sg/rieo/RI/Pages/Research-Data-Policies.aspx> (besucht am 21. 12. 2017).
- [Nat16] National Science Foundation. *Dear Colleague Letter: Data Management and Data Reporting Requirements for Research Awards Supported by the Office of Polar Programs*. 2016. URL: https://www.nsf.gov/pubs/2016/nsf16055/nsf16055.jsp?WT.mc_id=USNSF_25&WT.mc_ev=click (besucht am 21. 12. 2017).

- [nes16] nestor-Arbeitsgruppe Media, Hrsg. *Leitfaden für die digitale Langzeitarchivierung audiovisueller Medien*. Bd. 19. nestor-materialien. Berlin: nestor - Kompetenznetzwerk Langzeitarchivierung, 2016. URL: http://files.dnb.de/nestor/materialien/nestor_mat_19.pdf.
- [Neu+12] Heike Neuroth u. a., Hrsg. *Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme*. Boizenburg: Verlag Werner Hülsbusch, 2012. ISBN: 978-3-86488-008-7. URL: <http://nbn-resolving.de/urn:nbn:de:0008-2012031401>.
- [New17] Johanna Sophie Newell. "Ein internationaler Vergleich universitärer Forschungsdatenmanagement-Aktivitäten". Magisterarb. Berlin: Humboldt-Universität zu Berlin, 2017. DOI: 10.18452/18541. URL: <https://edoc.hu-berlin.de/handle/18452/19237#>.
- [nK17] nestor und Sabrina Kistner Hidalgo. *AG Media*. v. 39. 2017. URL: <https://wiki.dnb.de/display/NESTOR/AG+Media> (besucht am 19. 01. 2018).
- [Now17] Krzysztof Nowak. *Extended grant support, ORCID and language field*. Genf, 2017. URL: <http://blog.zenodo.org/2017/10/10/grants-and-language-update/> (besucht am 29. 12. 2017).
- [NS14] Lars Holm Nielsen und Tim Smith. *Zenodo Overview [Poster]*. Genf, 2014. DOI: 10.5281/zenodo.8428. URL: <https://doi.org/10.5281/zenodo.8428> (besucht am 29. 12. 2017).
- [NTU Ja] NTU Libraries. *Why share data?* Singapur, o. J. URL: <http://blogs.ntu.edu.sg/lib-datamanagement/whyshare/> (besucht am 21. 12. 2017).
- [NTU Jb] NTU Research Data Management Team. *About us*. Singapur, o. J. URL: <https://blogs.ntu.edu.sg/lib-datamanagement/contact-us/> (besucht am 21. 12. 2017).
- [NTU17a] NTU Libraries. *About DR-NTU (Data)*. Singapur, 2017. URL: <http://blogs.ntu.edu.sg/lib-datamanagement/about-dr-ntu-data/> (besucht am 21. 12. 2017).
- [NTU17b] NTU Libraries. *Data Exploration Guide*. Singapur, 2017. URL: https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/DRNTUData_DataExplorationGuide-1csermr.pdf (besucht am 21. 12. 2017).
- [NTU17c] NTU Libraries. *Dataset Metadata in DR-NTU (Data): Poster*. Singapur, 2017. URL: https://blogs.ntu.edu.sg/lib-scholarlycomm/files/2017/11/DRNTU_Poster7_v10-14bfaql.pdf (besucht am 21. 12. 2017).
- [NTU17d] NTU Libraries. *Dataset Records in DR-NTU (Data): Poster*. Singapur, 2017. URL: https://blogs.ntu.edu.sg/lib-scholarlycomm/files/2017/11/DRNTU_Poster6_v11-x2tsmr.pdf (besucht am 21. 12. 2017).
- [NTU17e] NTU Libraries. *Dataverse Community Norms*. Singapur, 2017. URL: <https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/Dataverse-Community-Norms-1f0fx4u.pdf> (besucht am 21. 12. 2017).
- [NTU17f] NTU Libraries. *Dataverse, Dataset and File Management*. Singapur, 2017. URL: https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/DRNTUData_DataverseDatasetandFileManagement-2kjeolg.pdf (besucht am 21. 12. 2017).
- [NTU17g] NTU Libraries. *Dataverse Software in DR-NTU (Data)*. Singapur, 2017. URL: https://blogs.ntu.edu.sg/lib-scholarlycomm/files/2017/11/DRNTU_Poster3_v11-1db143e.pdf (besucht am 21. 12. 2017).

- [NTU17h] NTU Libraries. *DR-NTU (Data) FAQ*. Singapur, 2017. URL: <http://blogs.ntu.edu.sg/lib-datamanagement/dr-ntu-data-faq/> (besucht am 21. 12. 2017).
- [NTU17i] NTU Libraries. *DR-NTU (Data) General Terms of Use*. Singapur, 2017. URL: https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/12/DRNTUData_GeneralTermsOfUse_v2_01122017-2azwrc.pdf (besucht am 21. 12. 2017).
- [NTU17j] NTU Libraries. *Finding and Using Data*. Singapur, 2017. URL: https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/DRNTUData_FindingUsingData-2mg2e94.pdf (besucht am 21. 12. 2017).
- [NTU17k] NTU Libraries. *How to upload*. Singapur, 2017. URL: https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/DRNTUData_HowToUpload-2bmpgd1.pdf (besucht am 21. 12. 2017).
- [NTU17l] NTU Libraries. *Supported Metadata and References*. Singapur, 2017. URL: https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/DRNTUData_SupportedMetadataReferences-2e09yn2.pdf (besucht am 21. 12. 2017).
- [NTU17m] NTU Libraries. *Tabular Data File Ingest*. Singapur, 2017. URL: https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/DRNTUData_TabularDataFileIngest-2k6g8bo.pdf (besucht am 21. 12. 2017).
- [NTU17n] NTU Libraries. *Where to share data?* Singapur, 2017. URL: <https://blogs.ntu.edu.sg/lib-datamanagement/wheretoshare/> (besucht am 21. 12. 2017).
- [Off12] Nicole Offhaus. "Institutionelle Repositorien und Universitätsbibliotheken: Entwicklungsstand und Perspektiven". Magisterarb. Köln: Fachhochschule Köln, 2012. URL: <http://fiz1.fh-potsdam.de/volltext/fhkoeln/12026.pdf>.
- [Ope15] Open Knowledge. *Open Definition - Version 2.1*. 2015. URL: <http://opendefinition.org/od/2.1/en/> (besucht am 15. 01. 2018).
- [Pam14] Heinz Pampel. *re3data.org - Stand und Perspektive*. 103. Deutscher Bibliothekartag, Bremen, 04.06.2014. 2014. URL: <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0290-opus-15756> (besucht am 21. 12. 2017).
- [PNH14] George Pyrounakis, Mara Nikolaidou und Michael Hatzopoulos. "Building Digital Collections Using Open Source Digital Repository Software". In: *International Journal of Digital Library Systems* 4.1 (2014), S. 10–24. ISSN: 1947-9077. DOI: [10.4018/ijdl.2014010102](https://doi.org/10.4018/ijdl.2014010102).
- [Rai15] Jean Rainwater. *What is ORCID?* Providence, RI, 2015. URL: <https://library.brown.edu/DigitalTechnologies/iiif-and-the-bdr/> (besucht am 31. 12. 2017).
- [Rai16] Roberto Raieli. "Introducing Multimedia Information Retrieval to libraries". In: *JLIS.it* 7.3 (2016), S. 9–42. ISSN: 2038-1026. DOI: [10.4403/jlis.it-11530](https://doi.org/10.4403/jlis.it-11530). URL: <https://www.jlis.it/article/view/11530>.
- [RB10] Jenn Riley und Devin Becker. *Seeing Standards: A Visualization of the Metadata Universe*. 2010. URL: <http://jennriley.com/metadatamap/> (besucht am 19. 01. 2018).
- [RDA16] RDA. *Matrix of use cases and functional requirements for research data repository platforms*. 2016. URL: <https://www.rd-alliance.org/group/repository-platforms-research-data-ig/outcomes/matrix-use-cases-and-functional-requirements>.

- [re316a] re3data.org. *Brown Digital Repository*. 2016. DOI: [10.17616/R3193B](https://doi.org/10.17616/R3193B). URL: <http://doi.org/10.17616/R3193B> (besucht am 21. 12. 2017).
- [re316b] re3data.org. *OceanRep GEOMAR Repository*. 2016. DOI: [10.17616/R3M927](https://doi.org/10.17616/R3M927). URL: <http://doi.org/10.17616/R3M927> (besucht am 21. 12. 2017).
- [re317a] re3data.org. *DR-NTU (Data)*. 2017. DOI: [10.17616/R3N77R](https://doi.org/10.17616/R3N77R). URL: <http://doi.org/10.17616/R3N77R> (besucht am 21. 12. 2017).
- [re317b] re3data.org. *LSE Research Online*. 2017. DOI: [10.17616/R3TC9F](https://doi.org/10.17616/R3TC9F). URL: <http://doi.org/10.17616/R3TC9F> (besucht am 21. 12. 2017).
- [re317c] re3data.org. *Marine Geoscience Data System*. 2017. DOI: [10.17616/R38C83](https://doi.org/10.17616/R38C83). URL: <http://doi.org/10.17616/R38C83> (besucht am 21. 12. 2017).
- [re317d] re3data.org. *Phaidra Universität Wien*. 2017. DOI: [10.17616/R3G606](https://doi.org/10.17616/R3G606). URL: <http://doi.org/10.17616/R3G606> (besucht am 21. 12. 2017).
- [re317e] re3data.org. *Stanford Digital Repository*. 2017. DOI: [10.17616/R3RK6R](https://doi.org/10.17616/R3RK6R). URL: <http://doi.org/10.17616/R3RK6R> (besucht am 21. 12. 2017).
- [re317f] re3data.org. *Zenodo*. 2017. DOI: [10.17616/R3QP53](https://doi.org/10.17616/R3QP53). URL: <http://doi.org/10.17616/R3QP53> (besucht am 21. 12. 2017).
- [RRK13] Anke Reinhold, Marc Rittberger und Stefan Keil. "A Thematic Analysis to Facilitate Data Curation Relevance Decisions: The case of Research Data in Video Studies". In: *Informationswissenschaft zwischen virtueller Infrastruktur und materiellen Lebenswelten*. Hrsg. von H. C. Hobohm. Glückstadt: Werner Hülsbusch, 2013, S. 198–209.
- [RS16] Robin Charlotte Rice und John Southall. *The data librarian's handbook*. London: Facet Publishing, 2016. ISBN: 9781783300471. URL: <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1454580>.
- [Rüc+15] Jessika Rücknagel u. a. *Metadata Schema for the Description of Research Data Repositories*. 2015. DOI: [10.2312/re3.008](https://doi.org/10.2312/re3.008).
- [Sch14] Barbara Schmidt. "Ein zweites Standbein für die Bibliothek OceanRep - das Institutional Repository des GEOMAR Helmholtz-Zentrums für Ozeanforschung Kiel". In: *BuB : Forum für Bibliothek und Information* 66.1 (2014), S. 57–60. URL: <http://oceanrep.geomar.de/id/eprint/23105>.
- [Sch16] Barbara Schmidt. *Eindeutige Autorenidentifikation*. FSR: Für Sie Recherchiert, 5. Kiel, 2016. DOI: [10.3289/FSR_5](https://doi.org/10.3289/FSR_5). (Besucht am 21. 12. 2017).
- [Sim12] Elena Simukovic. "Enhanced publications - Integration von Forschungsdaten beim wissenschaftlichen Publizieren". Magisterarb. Berlin: Humboldt-Universität zu Berlin, 2012. DOI: [10.18452/14182](https://doi.org/10.18452/14182). URL: <https://edoc.hu-berlin.de/handle/18452/14834#>.
- [Sny14] Stuart Snyderman. *Stanford begins development on Spotlight*. Providence, RI, 2014. URL: <https://library.stanford.edu/blogs/digital-library-blog/2014/02/stanford-begins-development-spotlight> (besucht am 21. 12. 2017).

- [Sny16] Stuart Snyder. *Stanford Libraries introduces new features in SearchWorks to support enhanced access to image collections*. Providence, RI, 2016. URL: <https://library.stanford.edu/blogs/digital-library-blog/2016/08/iiif-in-searchworks> (besucht am 26.12.2017).
- [Sta Ja] Stanford Library. *Advanced metadata*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/data-management-services/data-best-practices/creating-metadata/advanced-metadata> (besucht am 21.12.2017).
- [Sta Jb] Stanford Library. *Copyright, licenses and access restrictions*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/stanford-digital-repository/faqs/copyright-licenses-and-access-restrictions> (besucht am 21.12.2017).
- [Sta Jc] Stanford Library. *Data preservation at Stanford*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/data-management-services/share-and-preserve-research-data/data-preservation-stanford> (besucht am 21.12.2017).
- [Sta Jd] Stanford Library. *Depositor services*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/stanford-digital-repository/depositor-services> (besucht am 21.12.2017).
- [Sta Je] Stanford Library. *FAQs [SDR]*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/stanford-digital-repository/faqs> (besucht am 21.12.2017).
- [Sta Jf] Stanford Library. *Featured collections*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/stanford-digital-repository/sdr-benefits/featured-collections> (besucht am 21.12.2017).
- [Sta Jg] Stanford Library. *Pricing*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/stanford-digital-repository/depositor-services/pricing> (besucht am 21.12.2017).
- [Sta Jh] Stanford Library. *Repositories and data preservation FAQ*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/data-management-services/faq/repositories-and-data-preservation-faq> (besucht am 21.12.2017).
- [Sta Ji] Stanford Library. *SDR overview*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/stanford-digital-repository/sdr-overview> (besucht am 21.12.2017).
- [Sta Jj] Stanford Library. *SDR technology*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/stanford-digital-repository/sdr-overview/sdr-technology> (besucht am 21.12.2017).
- [Sta Jk] Stanford Library. *Technical topics*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/stanford-digital-repository/faqs/technical-topics> (besucht am 21.12.2017).
- [Sta Jl] Stanford Library. *When to use SDR*. Providence, RI, o. J. URL: <https://library.stanford.edu/research/stanford-digital-repository/sdr-benefits/when-use-sdr> (besucht am 21.12.2017).

- [Sta13] Stanford Library. *Digital collections now viewable in SearchWorks*. Providence, RI, 2013. URL: <https://library.stanford.edu/blogs/digital-library-blog/2013/03/digital-collections-now-viewable-searchworks> (besucht am 21. 12. 2017).
- [Sta17] Stanford Library. *Hydra-in-a-Box*. Providence, RI, 2017. URL: <https://library.stanford.edu/projects/hydra-in-a-box> (besucht am 26. 12. 2017).
- [Ste17] Christoph Steinhof. "Erfolgskriterien für Forschungsdatenrepositorien und deren Relevanz für verschiedene Stakeholder-Gruppen". Masterarbeit. Potsdam: Fachhochschule Potsdam, 2017. URL: <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:kobv:525-19521> (besucht am 16. 01. 2018).
- [UK oJ] UK Data Service. *Recommended formats*. o.J. URL: <https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats> (besucht am 21. 12. 2017).
- [Uni Ja] United Nations Office for Outer Space Affairs. *Marine Geoscience Data System (MGDS)*. o. J. URL: <http://www.un-spider.org/links-and-resources/institutions/marine-geoscience-data-system-mgds> (besucht am 30. 12. 2017).
- [Uni Jb] Universität Wien. *Phaidra*. Wien, o. J. URL: <http://phaidra.org/> (besucht am 22. 12. 2017).
- [Uni Jc] Universitätsbibliothek Wien. *Das Phaidra Netzwerk*. Wien, o. J. URL: <http://datamanagement.univie.ac.at/ueber-phaidra/das-phaidra-netzwerk/> (besucht am 22. 12. 2017).
- [Vie+14] Paul Vierkant u. a. *Schema for the description of research data repositories*. 2014. DOI: 10.2312/re3.006. URL: <http://gfzpublic.gfz-potsdam.de/pubman/item/escidoc:758898> (besucht am 05. 12. 2017).
- [VK14] Paul Vierkant und Maxi Kindling. "Welche Institutionen betreiben Open-Access-Repositorien in Deutschland?" In: *Libreas* 26 (2014). ISSN: 1860-7950. URL: <http://libreas.eu/ausgabe26/07vierkantkindling/>.
- [Voi14] Michaela Voigt. "Nutzungsrechte bei institutionellen Repositorien in Deutschland - Analyse der Rechteübertragung und Metadatenangaben". Magisterarb. Berlin: Humboldt-Universität zu Berlin, Philosophische Fakultät I, 2014. DOI: 10.18452/14233. URL: <https://edoc.hu-berlin.de/handle/18452/14885#> (besucht am 02. 01. 2018).
- [VW13] Sven Vlaeminck und Gert G. Wagner. *Zur Rolle von Forschungsdatenzentren beim Management von publikationsbezogenen Forschungsdaten: Ergebnisse einer Befragung von wissenschaftlichen Infrastrukturdienstleistern*. Working Paper Series des Rates für Sozial- und Wirtschaftsdaten. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD), 2013. URL: <http://hdl.handle.net/10419/84969>.
- [VW14] Sven Vlaeminck und Gert G. Wagner. "On the role of research data centres in the management of publication-related research data". In: *LIBER Quarterly* 23.4 (2014), S. 336–357. DOI: 10.18352/lq.9356.
- [Wen17] Chew Shu Wen. *DR-NTU (Data) is here!* Singapur, 2017. URL: <https://blogs.ntu.edu.sg/lib-scholarlycomm/?p=19030> (besucht am 21. 12. 2017).
- [Wil17] Helen K. R. Williams. "Ever evolving: metadata services and repository involvement at LSE". In: *Catalogue and Index* 187 (2017), S. 2–4. ISSN: 0008-7629. URL: <http://eprints.lse.ac.uk/id/eprint/82667> (besucht am 30. 12. 2017).

- [Wri12] Richard Wright. *Preserving Moving Pictures and Sound*. Techn. Ber. Digital-PreservationCoalition, 2012. DOI: [10.7207/twr12-01](https://doi.org/10.7207/twr12-01). URL: <http://www.dpconline.org/docs/technology-watch-reports/753-dpctw12-01-pdf/file>.
- [Zen Ja] Zenodo. *About Zenodo*. Genf, o. J. URL: <http://about.zenodo.org/> (besucht am 21. 12. 2017).
- [Zen Jb] Zenodo. *Features - Introducing Zenodo!* Genf, o. J. URL: <http://help.zenodo.org/features/> (besucht am 29. 12. 2017).
- [Zen Jc] Zenodo. *Frequently Asked Questions [Zenodo]*. Genf, o. J. URL: <http://help.zenodo.org/> (besucht am 21. 12. 2017).
- [Zen Jd] Zenodo. *General Policies*. Genf, o. J. URL: <http://about.zenodo.org/policies/> (besucht am 21. 12. 2017).
- [Zen Je] Zenodo. *Infrastructure*. Genf, o. J. URL: <http://about.zenodo.org/infrastructure/> (besucht am 29. 12. 2017).
- [Zen Jf] Zenodo. *New Upload*. Genf, o. J. URL: <https://zenodo.org/deposit/new> (besucht am 29. 12. 2017).
- [Zen Jg] Zenodo. *Principles*. Genf, o. J. URL: <http://about.zenodo.org/principles/> (besucht am 29. 12. 2017).
- [Zen Jh] Zenodo. *REST API*. Genf, o. J. URL: <http://developers.zenodo.org/#rest-api> (besucht am 29. 12. 2017).
- [Zen Ji] Zenodo. *Terms of Use v1.0*. Genf, o. J. URL: <http://about.zenodo.org/terms/> (besucht am 29. 12. 2017).
- [Zen17] Zenodo. *What's new in Zenodo?* Genf, 2017. URL: <http://help.zenodo.org/whatsnew/> (besucht am 29. 12. 2017).

Eigenständigkeitserklärung

Ich versichere hiermit, dass die vorliegende Bachelorarbeit mit dem Titel "Vergleich und Analyse von Forschungsdatenrepositorien" selbstständig verfasst und hierzu keine anderen als die angegeben Hilfsmittel verwendet habe.

Alle Stellen der Arbeit, die wörtlich oder sinngemäß aus fremden Quellen entnommen wurden, sind als solche kenntlich gemacht.

Die Arbeit wurde bisher an keiner anderen Stelle als Prüfungsleistung vorgelegt oder veröffentlicht.

Ich bin mir bewusst, dass eine falsche Erklärung rechtliche Folgen hat.

Berlin, den 29. Januar 2018



Melanie Janßen

A Anhänge

A.1. Analyse der Forschungsdatenrepositorien aus re3data

Liste der Repositorien

Alle 92 FDR, die nach Einschränkung über die Metadaten aus re3data extrahiert wurden, wurden in einer ersten Analyse untersucht. Eine laufende Nummer wurde zur Identifikation vergeben, die Reihenfolge der FDR entspricht der Standardanzeige in re3data und ist wertungsfrei. Die Spalte "Name" enthält den Namen des Repositoriums, auch hier wurden die Angaben aus re3data übernommen. Die Spalte "Link" wurde befüllt mit den URLs zu den Startseiten oder Sucheinstiegen der Repositorien. "Anzahl Suchergebnisse" bezieht sich auf alle Ergebnisse, die ein Repository am Stichtag anzeigte, die Ermittlung geschah durch eine leere Suche, wenn möglich. In Ausnahmefällen (z.B. für EPrints Plattformen) wurden alle Dokumentarten ausgewählt und die dort angezeigte Anzahl an Ergebnissen verwendet. Die Spalte "davon Video" enthält die Anzahl von Videos im FDR, die über Facettierung oder Erweiterte Suche ermittelt wurde. Akzeptiert wurden Facetten wie "audiovisuell", "Video", "moving image" etc. Der prozentuale Anteil wurde aus den beiden vorigen Spalten errechnet und sollte ursprünglich als Schwellenwert dienen, welches aufgrund der potentiell impliziten Wertung verworfen wurde. Die "Ausschluss-ID" ist eine laufende Nummer für die folgende Spalte und zur Vereinfachung einer automatisierten Zählung der folgenden Seite genutzt. Die Spalte "Ausschlussgrund" enthält acht Gründe, die FDR für die tiefere Analyse disqualifizierten:

- Entspricht nicht contentType Kriterium: Es konnten keine Videos im FDR festgestellt werden.
- Entspricht nicht dataAccess Kriterium: Der Zugang zu den Daten ist, auch nach Registrierung, nicht offen oder nicht über die Webseite möglich.
- Entspricht nicht repositoryLanguages Kriterium: Die Oberfläche der Webseite ist nicht in englisch oder deutsch. Auch wenn nur Hilfeseiten auf englisch oder deutsch vorliegen, aber die Schlagworte u.ä. nicht.
- keine Suche nach AV Daten: Eine Einschränkung auf Videos, audiovisuelle Daten oder Bewegtbilder wurde nicht ermöglicht. Beispielsweise wenn Datentyp oder Ressourcentyp nicht durchsuchbar sind oder keine Facettierung angeboten wird.
- keine passende Suchfunktion: FDR die gleichzeitig keine freie Textsuche und keine Facettierung anbieten. Beispielsweise statische Listen (Spezifika in Spalte "Details").
- FDR Offline: Das FDR existiert laut eigener Aussage nicht mehr. Wurde nicht für tote Links etc. verwendet, in denen das FDR durch Recherche gefunden werden konnte.

- kein FDR auf Webseite: Auf der Webseite und nach Recherche konnte auf den Webseiten kein öffentlich zugängliches FDR und kein Verweis auf ein solches gefunden werden (Spezifika in Spalte "Details").
- nur Portal, Datenlieferant separat betrachtet: Wurde verwendet für Webseiten die nur auf andere FDR verweisen, Metasuchen für FDR oder Mehrfacheinträge und Dopplungen in re3data. Gilt nicht für FDR die mehrere getrennte Sucheinstiege anbieten, aber eine gemeinsame Infrastruktur und Richtlinien haben.

Die Spalte "Details" enthält Spezifika des jeweiligen FDR, die durch den Ausschlussgrund allein nicht deutlich wurden; meist handelt es sich um weitere Angaben. In einem Fall wurde eine Inhaltswarnung ausgesprochen. Die Spalte "Screenshot" enthält die Dateinamen der Screenshots, die die getroffenen Aussagen belegen. Screenshots zeigen vor allem die Anzahl der Ergebnisse, Sucheinstiege, Suchmöglichkeiten etc. Die Benennung erfolgt nach dem Schema laufendeNummer_Name_[Beschreibung der Aussage], wobei letzteres nur für multiple Screenshots pro FDR verwendet wurde. Die Screenshots liegen der Arbeit im Ordner "Erste Analyse der FDR aus re3data" auf Datenträger bei.

lfd. Nr.	Name	Link	Anzahl Suchergebnisse	davon Video	Prozent. Anteil	Ausschluss-ID	Ausschlussgrund	Details	Screenshot
1	CLARIN-D Centre Leipzig ASV Automatische Sprachverarbeitung	http://clarin.informatik.uni-leipzig.de/repo/	7,473	0		1	Entspricht nicht content-Types Kriterium	Suche über VLO oder Fedora Commons Search Interface, keine Videodaten (vgl. Portal CLARIN-ERIC #24)	001_CLARIN-D Centre Leipzig ASV Automatische Sprachverarbeitung_Fedora 001_CLARIN-D Centre Leipzig ASV Automatische Sprachverarbeitung_VLO 002_GAMS
2	GAMS	http://gams.uni-graz.at/context:gams	38,430			4	keine Suche nach AV Daten		003_Hamburger Zentrum für Sprachkorpora Korpus Repository
3	Hamburger Zentrum für Sprachkorpora Korpus Repository	https://corpora.uni-hamburg.de/hzsk/en/repository-search	53			4	keine Suche nach AV Daten		004_OceanRep GEOMAR Repository_Suche 004_OceanRep GEOMAR Repository_Ergebnis
4	OceanRep GEOMAR Repository	http://oceanrep.geomar.de/	21,435	28	0.1%				005_Phaidra Universität Wien
5	Phaidra Universität Wien	https://phaidra.univie.ac.at/	71,544	1,146	2%				006_Qualiservice
6	Qualiservice	http://www.qualiservice.org/index.php?id=44	n.a.			2	Entspricht nicht dataAccess Kriterium	Daten nur per Zusendung auf Anfrage, keine Suchfunktion	007_Research Data Centre for Education_Aufzeichnung 007_Research Data Centre for Education_Studie
7	Research Data Centre for Education	https://www.fdz-bildung.de/datenarchiv.php	1,241			4	keine Suche nach AV Daten	Ermittlung ungefähr möglich über Namen der Studien, aber keine explizite Suche nach Videos	008_UdS Fedora Commons Repository_Fedora 008_UdS Fedora Commons Repository_VLO
8	UdS Fedora Commons Repository	http://fedora.clarin-d.uni-saarland.de/users.en.html	113	0		1	Entspricht nicht content-Types Kriterium	Suche über VLO oder Fedora Commons Search Interface, keine Videodaten (vgl. Portal CLARIN-ERIC #24)	009_Verbund Forschungsdaten Bildung
9	Verbund Forschungsdaten Bildung	https://www.forschungsdaten-bildung.de/studienliste.php?la=de	n.a.			4	keine Suche nach AV Daten	verzeichnet auch Daten aus Research Data Centre For Education (#7)	010_Kujawsko-Pomorska Digital Library
10	Kujawsko-Pomorska Digital Library	http://kpbk.umk.pl/dlibra	n.a.			3	Entspricht nicht repositoryLanguages Kriterium	Primär polnisch, vor allem Schlagworte etc.	011_Precipitation Measurement Missions_GPM 011_Precipitation Measurement Missions_TRMM
11	Precipitation Measurement Missions	https://pmm.nasa.gov/data-access	n.a.			5	keine passende Suchfunktion	Daten aus einer festen Anzahl an NASA Missionen, keine Suchfunktion sondern Auswahlliste	
12	UK Data Service	https://discover.ukdataservice.ac.uk/	7,355			8	nur Portal, Datenlieferant separat betrachtet	Datenlieferant: UK Data Archive (#18)	

Fortsetzung auf der nächsten Seite ...

...Fortsetzung der Liste der Repositorien

lfd. Nr.	Name	Link	Anzahl Suchergebnisse	davon Video	Prozent. Anteil	Ausschluss -ID	Ausschlussgrund	Details	Screenshot*
13	Plant Metabolic Network	https://www.plantcyc.org/	n.a.			4	keine Suche nach AV Daten		013_Plant Metabolic Network_EinfacheSuche 013_Plant Metabolic Network_ErweiterteSuche 014_Research Data Australia
14	Research Data Australia	https://researchdata.ands.org.au/	138,160			4	keine Suche nach AV Daten		
15	Partnership for Interdisciplinary Studies of Coastal Oceans	http://www.piscoweb.org/data	19,820			4	keine Suche nach AV Daten	Suche über DataOne (https://search.dataone.org/#data), Data source: PISCO MN	015_Partnership for Interdisciplinary Studies of Coastal Oceans
16	RIT Digital Media Library Repository	https://digitalarchive.rit.edu/xmlui/	11,123			4	keine Suche nach AV Daten		016_RIT Digital Media Library Repository
17	IMS Universität Stuttgart Repository	http://clarin04.ims.uni-stuttgart.de/fedora/objects	n.a.			4	keine Suche nach AV Daten	Suche nur über Fedora	017_IMS Universität Stuttgart Repository
18	UK Data Archive	http://www.data-archive.ac.uk/	7,355	32	0.4%			Durchsuchbar über UK Data Service (#12)	018_UK Data Archive 018_UK Data Archive_Ergebnis 018_UK Data Archive_Suche
19	The Cancer Imaging Archive	http://www.cancerimagingarchive.net/	74			4	keine Suche nach AV Daten		019_The Cancer Imaging Archive_EinfacheSuche 019_The Cancer Imaging Archive_ErweiterteSuche
20	The Cambridge Structural Database	https://www.ccdc.cam.ac.uk/solutions/csd-system/components/csd/	n.a.			4	keine Suche nach AV Daten	Erweiterte Suche nur mit Benutzerkonto (kostenlos) und Registrierung der Software (Kosten)	020_The Cambridge Structural Database
21	CSIRO Data Access Portal	https://data.csiro.au/dap/home?execution=e1s1	2,041			4	keine Suche nach AV Daten	Auch nicht für einzelne Domänen in Domain Search	021_CSIRO Data Access Portal_Browse 021_CSIRO Data Access Portal_DomainSearch 021_CSIRO Data Access Portal_Suche
22	National Science Digital Library	https://nsdl.oercommons.org/	46,983	11,518	25%				022_The National Science Digital Library_Ergebnisse 022_The National Science Digital Library_ErweiterteSuche 022_The National Science Digital Library_Suche
23	The Digital Archaeological Record	https://www.tdar.org/	418,374	0		1	Entspricht nicht contentTypes Kriterium	Nur Bilder etc.	023_The Digital Archaeological Record_Ergebnisse 023_The Digital Archaeological Record_Suche
24	CLARIN-ERIC	https://vlo.clarin.eu/?i	1,596,458			8	nur Portal, Datenlieferant(en) separat betrachtet	Datenlieferanten u.a. ASV (#1) & UdS (#8) Im Portal 32,539 Videodaten, 2% Anteil video/avi (457) video/mp4 (2748) video/mpeg (110) video/quicktime (3212) video/x-motion-jpeg (27) video/x-mpeg1 (25234) video/x-mpeg2 (13023) video/x-ms-wmv (1) 83 / 3%	024_CLARIN-ERIC_ErgebnisseGesamt 024_CLARIN-ERIC_ErgebnisseVideo 024_CLARIN-ERIC_Suche
25	META-SHARE	http://metashare.elda.org/repository/search/?q=	2,829			8	nur Portal, Datenlieferant (u.a. Center of Estonian Language Resources (#79) separat betrachtet		025_META-SHARE_ErgebnisseVideo 025_META-SHARE_Suche
26	LINDAT/CLARIN repository	https://lindat.mff.cuni.cz/repository/xmlui/	1,012			4	keine Suche nach AV Daten	Ausschnitt auch über VLO durchsuchbar (230 Ergebnisse, https://vlo.clarin.eu/Collection:LINDAT/CLARINData & Tools), dort: Entspricht nicht contentTypes Kriterium	026_LINDAT-CLARIN repository 026_LINDAT-CLARIN repository_VLO
27	Collaborative Research Centre Transregio 32 Database	http://www.tr32db.uni-koeln.de/site/index.php	n.a.			4	keine Suche nach AV Daten		027_Collaborative Research Centre Transregio 32 Database

Fortsetzung auf der nächsten Seite ...

...Fortsetzung der Liste der Repositorien

lfd. Nr.	Name	Link	Anzahl Such- ergebnisse	davon Video	Prozent Anteil	Ausschluss -ID	Ausschlussgrund	Details	Screenshot*
28	AVISO+	https://www.aviso.altimetry.fr/en/data/data-access.html	n.a.			5	keine passende Suchfunktion	keine freie Suche oder Browsing sondern Auswahlliste o.ä. bzw Zugang zur Suche erst nach Registrierung / Software-Download	028_AVISO+ 028_AVISO+_Suche
29	biodiversity.aq	http://data.biodiversity.aq/search_engine/es_search	47,181			4	keine Suche nach AV Daten		029_biodiversity.aq_Ergebnisse
30	Brown Digital Repository	https://repository.library.brown.edu/studio/search/?utf8=%E2%9C%93&search_field=q	181,224	104	0.1%				030_Brown Digital Repository_ErgebnisseGesamt 030_Brown Digital Repository_ErgebnisseVideo
31	Aston Data Explorer	http://researchdata.aston.ac.uk/cgi/search/advanced	154	0		1	Entspricht nicht content-Types Kriterium		031_Aston Data Explorer_ErgebnisseGesamt 031_Aston Data Explorer_ErgebnisseVideo 031_Aston Data Explorer_Suche
32	EDINA	https://edina.ac.uk/	n.a.			7	kein FDR auf Webseite	EDINA ist kein FDR, sondern Betreuer von Edinburgh DataShare (nicht in Liste wegen DataAccessRestriction Problem, siehe Methode)	032_EDINA
33	Antarctic & Southern Ocean Data Portal	http://www.marine-geo.org/tools/new_search/index.php?initiative=USAP	2,293			7	kein FDR auf Webseite	Kein eigenständiges FDR, nur eigene Suchmaske, sog. "Community Data Portal". FDR ist #54 Marine Geoscience Data System	033_Antarctic & Southern Ocean Data Portal_ErgebnisseGesamt 033_Antarctic & Southern Ocean Data Portal_ErgebnisseVideo 033_Antarctic & Southern Ocean Data Portal_Suche
34	Surface management system database	About">http://brainvis.wustl.edu/wiki/index.php/Sums>About				6	FDR Offline	Offline seit 2016	034_Surface management system database
35	National Archive of Computerized Data on Aging	http://www.icpsr.umich.edu/icpsrweb/NACDA/	1,527			4	keine Suche nach AV Daten		035_National Archive of Computerized Data on Aging
36	Henry A. Murray Research Archive	https://murray.harvard.edu/ https://dataverse.harvard.edu/dataverse/mra	6,244			2	Entspricht nicht dataAccess Kriterium	Aussage auf Webseite: "Our collection comprises over 100 [125] terabytes of data, audio, and video", aber keine Videodateien im FDR verfügbar, auch nicht in Kategorie "Studies with Video Data"	036_Henry A. Murray Research Archive_Ergebnisse 036_Henry A. Murray Research Archive_Suche
37	MultiDark Database	http://www.multidark.org/	n.a.			5	keine passende Suchfunktion	keine freie Suche oder Browsing sondern Auswahlliste o.ä. Eingebunden in CosmoSim (#88)	037_MultiDark Database_DatenbankenÜbersicht 037_MultiDark Database_Suche
38	Neuroimaging Informatics Tools and Resources Clearinghouse	https://www.nitrc.org/	996			4	keine Suche nach AV Daten		038_Neuroimaging Informatics Tools and Resources Clearinghouse_ErgebnisseGesamt 038_Neuroimaging Informatics Tools and Resources Clearinghouse_Suche
39	Macaulay Library of Natural Sounds	https://www.macaulaylibrary.org/	5,371,764			8	nur Portal, Datenlieferant separat betrachtet	Datenlieferant vorrangig eBird (#65), Im Portal 56,740 Videodaten, 1% Anteil	039_Macaulay Library of Natural Sounds_ErgebnisseGesamt 039_Macaulay Library of Natural Sounds_ErgebnisseVideo 039_Macaulay Library of Natural Sounds_Portal
40	Social Science Data Archive - UCLA	https://dataverse.harvard.edu/dataverse/ssda_ucla	4,358			4	keine Suche nach AV Daten		040_Social Science Data Archive - UCLA_Ergebnisse 040_Social Science Data Archive - UCLA_Suche
41	CESSDA ERIC	https://www.cessda.eu/	n.a.			7	kein FDR auf Webseite	Nur Koordination / Service / Infrastrukturangebote, bisher kein eigenes FDR	041_CESSDA ERIC
42	DR-NTU Data repository	https://researchdata.ntu.edu.sg/	162	1	1%				042_DR-NTU-ErgebnisseGesamt 042_DR-NTU-ErgebnisseVideo

Fortsetzung auf der nächsten Seite ...

...Fortsetzung der Liste der Repositorien

lfd. Nr.	Name	Link	Anzahl Suchergebnisse	davon Video	Prozent. Anteil	Ausschluss -ID	Ausschlussgrund	Details	Screenshot*
43	LSE Research Online	http://eprints.lse.ac.uk/	68,379	40	0.1%				043_London School of Economics and Political Science Research Online_ErgebnisseGesamt 043_London School of Economics and Political Science Research Online_ErgebnisseVideo 043_London School of Economics and Political Science Research Online_Suche
44	EASY	https://easy.dans.knaw.nl/ui/browse	41,519			4	keine Suche nach AV Daten		044_EASY_Browse 044_EASY_ErgebnisseGesamt 044_EASY_ErweiterteSuche
45	Ocean Networks Canada	http://dmas.uvic.ca/DataSearch	n.a.			5	keine passende Suchfunktion	kontinuierliche Videoaufnahme, Suche nur in Zeitintervallen	045_Ocean Networks Canada
46	Illinois Digital Environment for Access to Learning and Scholarship	https://www.ideals.illinois.edu/	94,025			4	keine Suche nach AV Daten	Suchfeld "Type" ist Freitext, kein Vokabular / Autovervollständigung / Datentypen hinterlegt	046_IDEALS_ErgebnisseGesamt 046_IDEALS_Suche
47	DIGITAL.CSIC	https://digital.csic.es/?locale=en	150,145			3	Entspricht nicht repositoryLanguages Kriterium	Primär spanisch, zusätzlich keine Suche nach AV Daten	047_DIGITAL.CSIC_Browse 047_DIGITAL.CSIC_Suche
48	University of Reading Research Data Archive	http://researchdata.reading.ac.uk/cgi/search/advanced	46	0		1	Entspricht nicht content-Types Kriterium		048_University of Reading Research Data Archive_ErgebnisseGesamt 048_University of Reading Research Data Archive_ErgebnisseVideo 048_University of Reading Research Data Archive_Suche
49	Eumetsat	http://navigator.eumetsat.int/discovery/Start/Explore/Quick.do	650			4	keine Suche nach AV Daten		049_Eumetsat_ErgebnisseGesamt 049_Eumetsat_Suche
50	CLARIN i Danmark	https://clarin.dk/clarindk/find.jsp	157,607			3	Entspricht nicht repositoryLanguages Kriterium	Primär Dänisch, enthält 80 Videodaten, 0,1% Anteil	050_CLARIN i Danmark_ErgebnisseGesamt 050_CLARIN i Danmark_ErgebnisseVideo 050_CLARIN i Danmark_Sprache 050_CLARIN i Danmark_Suche
51	GeoPRISMS Data Portal	http://www.marine-geo.org/tools/new_search/index.php?initiative=GeoPRISMS	515	0		1	Entspricht nicht content-Types Kriterium	Photograph:Video ist nur übergeordnet (alle Datensätze geprüft) und Photograph:Video:FrameGrab ist nur Bild. Zusätzlich: Kein eigenständiges FDR, nur sog. "Community Data Portal". FDR ist #54 Marine Geoscience Data System	051_GeoPRISMS Data Portal_ErgebnisseGesamt 051_GeoPRISMS Data Portal_ErgebnisseVideo 051_GeoPRISMS Data Portal_Suche
52	Centers for Disease Control and Prevention, Data & Statistics	https://data.cdc.gov/browse	659			4	keine Suche nach AV Daten		052_Centers for Disease Control and Prevention, Data & Statistics
53	California Vectorborne Disease Surveillance System	http://www.calsurv.org/taxonomy_menu/6	n.a.			4	keine Suche nach AV Daten		053_California Vectorborne Disease Surveillance System
54	Marine Geoscience Data System	http://www.marine-geo.org/tools/new_search/search_map.php	8,624	3	0.03%	8		Datenportale (äquivalent zu Sammlungen): Antarctic & Southern Ocean Data Portal (#33) und GeoPRISMS Data Portal (#51)	054_Marine Geoscience Data System_ErgebnisseGesamt 054_Marine Geoscience Data System_ErgebnisseVideo
55	NatureServe	http://explorer.natureserve.org/servlet/NatureServe#criteria	80,871			4	keine Suche nach AV Daten		055_NatureServe_ErgebnisseGesamt 055_NatureServe_Suche
56	The Global Agricultural Trial Repository and Database	http://www.agtrials.org/searchtrials	n.a.			4	keine Suche nach AV Daten		056_The Global Agricultural Trial Repository and Database_Suche

Fortsetzung auf der nächsten Seite ...

...Fortsetzung der Liste der Repositorien

lfd. Nr.	Name	Link	Anzahl Suchergebnisse	davon Video	Prozent. Anteil	Ausschluss-ID	Ausschlussgrund	Details	Screenshot*
57	Alaska Satellite Facility SAR Data Center	https://vertex.daac.asf.alaska.edu/	n.a.			4	keine Suche nach AV Daten		057_Alaska Satellite Facility SAR Data Center_Suche
58	Incorporated Research Institutions for Seismology	http://ds.iris.edu/ds/nodes/dmc/	n.a.			4	keine Suche nach AV Daten	Diverse Sucheinstiege, keine der Stichproben ermöglicht Suche nach AV Daten	058_Incorporated Research Institutions for Seismology_Suche1 058_Incorporated Research Institutions for Seismology_Suche2 058_Incorporated Research Institutions for Seismology_Suche3
59	Internet Archive	https://archive.org/search.php?query=%2A	31,665,910	3,773,506	12%				059_Internet Archive_ErgebnisseGesamt 059_Internet Archive_ErgebnisseVideo
60	Speech and Language Data Repository	http://www.sldr.org/				6	FDR Offline	Offline seit 2015, Nachfolger Ortolang (nicht in Liste wegen DataAccessRestriction Problem, re3data Datensatz http://doi.org/10.17616/R3T054)	060_Speech and Language Data Repository
61	ADA Dataverse	https://dataverse.ada.edu.au/	387			4	keine Suche nach AV Daten		061_ADA Dataverse
62	Koç University Digital Collections	http://digitalcollections.library.ku.edu.tr/cdm/	52,923			4	keine Suche nach AV Daten	Suchfeld "Format" ist Freitext, kein Vokabular / Autovervollständigung / Datentypen hinterlegt	062_Koc University Digital Collections_ErgebnisseGesamt 062_Koc University Digital Collections_Suche
63	National Center for Education Statistics	https://nces.ed.gov/surveys/international/ide/	n.a.			4	keine Suche nach AV Daten	Fünf Sucheinstiege, keiner erlaubt Suche nach AV Daten	063_National Center for Education Statistics 063_National Center for Education Statistics_Suche
64	HJ Andrews Experimental Forest	http://andlter.forestry.oregonstate.edu/data/catalog/datacatalog.aspx	181			4	keine Suche nach AV Daten		064_HJ Andrews Experimental Forest_Browse 064_HJ Andrews Experimental Forest_ErweiterteSuche
65	eBird	https://ebird.org/media/catalog	5,000,343	49,203	1%			Suchinterface von und Datenlieferant für Macaulay Library of Natural Sounds (#39)	065_eBird_Suche
66	DataBox	https://data.kb.dk/dvn/	456			4	keine Suche nach AV Daten	Suchfeld "Kind of Data" ist Freitext, kein Vokabular / Autovervollständigung / Datentypen hinterlegt	066_DataBox_ErgebnisseGesamt 066_DataBox_Suche
67	CLARIN-PL	https://clarin-pl.eu/dspace/discover?advance	280			4	keine Suche nach AV Daten		067_CLARIN-PL_Suche
68	Invasive.org	https://www.invasive.org/	258,238			5	keine passende Suchfunktion	nur Browsing in neusten Videos	068_Invasive.org_Browse
69	Global Multi-Resolution Topography Data Portal	http://media.marine-geo.org/album/global-multi-resolution-topography	n.a.			1	Entspricht nicht content-Types Kriterium	nur noch Previews der Videos, die Videodateien sind nicht mehr erreichbar	069_Global Multi-Resolution Topography Data Portal_Browse 069_Global Multi-Resolution Topography Data Portal_Videoanzeige
70	UCD Digital Library	https://digital.ucd.ie/?q=&start=1	156,305	298	0.2%			Resource type: moving image	070_UCD Digital Library_ErgebnisseGesamt 070_UCD Digital Library_ErgebnisseVideo
71	CINES Preservation	https://www.cines.fr/en/long-term-preservation/typologies/donnees-scientifiques/	n.a.			3	Entspricht nicht repositoryLanguages Kriterium		071_CINES-PAC
72	National Addiction & HIV Data Archive Program	http://www.icpsr.umich.edu/icpsrweb/NAHDAP/studies?archive=NAHDAP	417			4	keine Suche nach AV Daten		072_National Addiction & HIV Data Archive Program
73	World Data Centre for Glaciology	https://www.spri.cam.ac.uk/library/catalogue/sprilib/icesnow/	n.a.			4	keine Suche nach AV Daten		073_World Data Centre for Glaciology

Fortsetzung auf der nächsten Seite ...

...Fortsetzung der Liste der Repositorien

lfd. Nr.	Name	Link	Anzahl Suchergebnisse	davon Video	Prozent. Anteil	Ausschluss -ID	Ausschlussgrund	Details	Screenshot*
74	Constrained Local Universe Simulations	https://www.clues-project.org/cms/images-and-movies/dark-matter/	n.a.			5	keine passende Suchfunktion	nur Browsing, Eingebunden in CosmoSim (#88)	074_Constrained Local Universe Simulations_Browse
75	Space Physics Interactive Data Resource	http://spidr.ionosonde.net/spidr/dataset.do?task=clearview https://spidr.ngdc.noaa.gov/ http://dla.who.edu/catalog	n.a.			6	FDR Offline	Stand 21.12.17, Fehler für beide Zugänge (ERR_CONNECTION_RESET / ERR_NAME_NOT_RESOLVED)	075_Space Physics Interactive Data Resource_Offline
76	Woods Hole Oceanographic Institution Data Library and Archives	http://dla.who.edu/catalog	10,546			4	keine Suche nach AV Daten		0076_Woods Hole Oceanographic Institution Data Library and Archives_Suche
77	The Language Archive	https://corpus1.mpi.nl/ds/asv/?jsessionid=E5F7FD3E009CF39E88083F6A41FD95D7?0	n.a.			5	keine passende Suchfunktion	Suche nur in einzelnen Beständen, keine allgemeine / übergeordnete Suche, Fehlermeldungen bei Auswahl von "Format"	077_The Language Archive_Fehlermeldung
78	Stanford Digital Repository	https://searchworks.stanford.edu/?f%5Bbuilding_facet%5D%5B%5D=Stanford+Digital+Repository	259,422	4,536	2%				078_Stanford Digital Repository_ErgebnisseGesamt
79	Center of Estonian Language Resources	https://keeleressursid.ee/en/resources/speech-databases https://metashare.ut.ee/ https://zenodo.org/search?page=1&size=20&q=	123			3	Entspricht nicht repositoryLanguages Kriterium	Primär Estnisch, enthält 1 Videodatei, 1% Anteil	079_Center of Estonian Language Resources_ErgebnisseGesamt
80	Zenodo	https://zenodo.org/search?page=1&size=20&q=	311,103	442	0.1%				080_Zenodo_ErgebnisseGesamt
81	CLARIN.SI repository	https://www.clarin.si/repository/xmlui/	79			4	keine Suche nach AV Daten		081_CLARIN.SI repository
82	Sound and Vision / Beelden Geluid	http://in.beeldengeluid.nl/collectie/search?q=&publiclyViewableResultsOnly=true http://www.eyemoviepedia.com/search	42,217			3	Entspricht nicht repositoryLanguages Kriterium	Primär niederländisch, enthält 23,291 Videodateien, 55% Anteil	082_Sound and Vision_ErgebnisseGesamt
83	eyeMoviePedia	http://www.eyemoviepedia.com/search	n.a.			5	keine passende Suchfunktion	nur Browsing Warnung: sehr blutige Vorschaubilder (Operationsvideos am menschlichen Auge), Screenshot enthält keine Vorschaubilder	083_eyeMoviePedia
84	CLARIN repository at the University of Tübingen	https://vlo.clarin.eu/search?1	408			1	Entspricht nicht contentTypes Kriterium	Recherche über VLO (Collection: Tübingen Language Resources)	084_CLARIN repository at the University of Tübingen_ErgebnisseGesamt
85	Virgo Millenium Database	http://gavo.mpa-garching.mpg.de/Millennium/	n.a.			5	keine passende Suchfunktion	nur SQL Query	085_Virgo Millenium Database_Suche
86	VENUS in the Salish Sea	http://www.oceannetworks.ca/installations/observatories/venus-salish-sea https://phenome.jax.org/	n.a.			7	kein FDR auf Webseite	re3data Link nicht länger funktional, kein FDR auf Webseite http://www.oceannetworks.ca/ gefunden	086_VENUS in the Salish Sea
87	Mouse Phenome Database	https://phenome.jax.org/	n.a.			5	keine passende Suchfunktion	nur Browsing	087_Mouse Phenome Database_Suche
88	CosmoSim	https://www.cosmosim.org/	n.a.			8	nur Portal, Datenlieferant separat betrachtet	Datenlieferant MultiDark Database (#37) und Constrained Local Universe Simulations (#74), zusätzlich keine passende Suchfunktion, nur SQL Query / Skript	088_CosmoSim

Fortsetzung auf der nächsten Seite ...

...Fortsetzung der Liste der Repositorien

lfd. Nr.	Name	Link	Anzahl Suchergebnisse	davon Video	Prozent. Anteil	Ausschluss-ID	Ausschlussgrund	Details	Screenshot*
89	Meertens Instituut Collecties	http://www.meertens.knaw.nl/cms/en/collections	n.a.			7	kein FDR auf Webseite	diverse Kollektionen und Datenbanken, keine übergreifende Suche nach Forschungsdaten	089_Meertens Instituut Collecties
90	York Digital Library	https://dlib.york.ac.uk/yodl/app/home/search?query=*	106,006			4	keine Suche nach AV Daten	Kein passender Eintrag in "Type" oder "Format"	090_York Digital Library_ErgebnisGesamt 090_York Digital Library_Suche
91	heiDATA	https://heidata.uni-heidelberg.de/	631			1	Entspricht nicht content-Types Kriterium		091_heidata_ErgebnisseGesamt 091_heidata_ErgebnisseSuche
92	United States Census Bureau	https://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t	n.a.			4	keine Suche nach AV Daten		092_United States Census Bureau_Browse 092_United StatesCensusBureau_Suche

Tabelle A.1.: Liste der Repositorien aus re3data

* Die Screenshots aus der Tabelle A.1 liegen der Arbeit im Ordner "Erste Analyse der FDR aus re3data" auf CD-ROM bei.

Ergebnisse der ersten Analyse

In der Spalte "FDR" sind die kriterienkonformen Repositorien aufgelistet, die Sortierung erfolgte nach der Anzahl der audiovisuellen Datensätze in der folgenden Spalte.

FDR	Anzahl an AV- Ergebnissen
Internet Archive	3,773,506
eBird	49,203
National Science Digital Library	11,518
Stanford Digital Repository	4,536
Phaidra Universität Wien	1,146
Zenodo	442
UCD Digital Library	298
Brown Digital Repository	104
LSE Research Online	40
UK Data Archive	32
OceanRep GEOMAR Repository	28
Marine Geoscience Data System	3
DR-NTU Data repository	1

Tabelle A.2.: Ergebnisse der ersten Analyse: Anzahl an AV

In der Spalte "FDR" sind die kriterienkonformen Repositorien aufgelistet, die Sortierung erfolgte nach dem prozentualen Anteil von audiovisuellen Forschungsdaten in der folgenden Spalte.

FDR	Prozentualer Anteil an AV
National Science Digital Library	25%
Internet Archive	12%
Stanford Digital Repository	2%
Phaidra Universität Wien	2%
eBird	1%
DR-NTU Data repository	1%
UK Data Archive	0.4%
UCD Digital Library	0.2%
Zenodo	0.1%
OceanRep GEOMAR Repository	0.1%
LSE Research Online	0.1%
Brown Digital Repository	0.1%
Marine Geoscience Data System	0.03%

Tabelle A.3.: Ergebnisse der ersten Analyse: Prozentualer Anteil an Videos im Vergleich zu allen Ergebnissen

In der Spalte "FDR" sind die kriterienkonformen Repositorien aufgelistet, die Sortierung erfolgte nach der Anzahl der AV-Ergebnisse. Die Spalte Art der Datenaggregation zeigt das Ergebnis der Prüfung der Datenaggregation. Push weist in diesem Kontext auf eine automatische Aggregation hin, also der Übernahme von Daten aus anderen Repositorien. Als Pull markierte Repositorien erhalten ihre Forschungsdaten durch direkte Eingabe von Personen, entweder Forscherx oder Mitarbeiterx des FDR. Repositorien die beide Varianten erlauben sind als Push-Pull markiert. Die Screenshots der Webseiten zeigen den Beleg für die Art der Datenaggregation.

FDR	Art der Datenaggregation	Screenshots der Webseiten**
Internet Archive	Push & Pull	Internet Archive_Push-Pull
eBird	Push & Pull	eBird_Push-Pull
National Science Digital Library	Push	The National Science Digital Library_Push
Stanford Digital Repository	Pull	Stanford Digital Repository_Pull
Phaidra Universität Wien	Pull	Phaidra Universität Wien_Pull
Zenodo	Pull	Zenodo_Pull
UCD Digital Library	Push & Pull	UCD Digital Library_Push-Pull
Brown Digital Repository	Pull	Brown Digital Repository_Push-Pull
LSE Research Online	Pull	London School of Economics and Political Science Research Online_Pull
UK Data Archive	Push & Pull	UK Data Archive_Push-Pull
OceanRep GEOMAR Repository	Pull	OceanRep GEOMAR Repository_Pull
Marine Geoscience Data System	Pull	Marine Geoscience Data System_Pull
DR-NTU Data repository	Pull	DR-NTU Data repository - Nanyang Technological University_Pull

Tabelle A.4.: Ergebnisse der ersten Analyse: Push/Pull Prüfung

** Die Screenshots aus der Tabelle A.4 liegen der Arbeit im Ordner "Analyse der kriterienkonformen FDR > Push-Pull Prüfung" auf CD-ROM bei.

A.2. Ausgefüllte Kriterienkataloge der FDR

Die Kriterien auf Level 1 und 2 wurden direkt von [Ste17] übernommen, die Unterpunkte auf Level 3 wurden aus dem Fließtext extrahiert. Wenn möglich wurden direkte Zitate von den Webseiten übernommen, die als Grundlage für Kapitel 5 dienen.

K1 Veröffentlichung und Nachnutzung

K1.1 Datenabgabe

- Zulässige Dateiformate
- Möglichkeiten des Uploads
- Datenintegritätsprüfung

K1.3 Dokumentation

- Dokumentation der FD
- Link zu Datenpublikation

K1.5 Lizenzmodelle

- Nutzungsbedingungen
- Verwendung von offenen Lizenzen
- Datenschutz
- MD Bindungslos lizenziert

K1.7 Nutzungsstatistik

- Statistik wird Erhoben
- Anzeige Statistik auf Datensatzebene
- standardisierte Erhebung

K2 Zugriff und Auffindbarkeit

K2.1 Zugangsformen

- Zugangsbeschränkung
- Nutzerkreis der Daten abgeben kann
- Zugriffsteuerung für sensible Daten
- Metadaten immer sichtbar

K2.3 Recherchemöglichkeiten

- Freie Textsuche
- Suchoperatoren
- Attributspezifische Suche
- Facettierung
- Empfehlungssysteme
- Volltextsuche
- Indexierbare Landingpage

K2.5 Programmierschnittstellen

K1.2 Metadaten

- Metadatenkategorien
- standardisierte Metadatenformate
- unterstützende Tools für die Eingabe von MD
- formularbasierte Eingabe von MD
- Hinweis auf Pflichtangaben
- Hilfe bei Schlagworten

K1.4 Zitierfähigkeit

- Zitationsrichtlinie
- Persistent Identifier

K1.6 Verweise auf Veröffentlichungen

- Link zu Publikation die auf FD basiert
- Link zu nachnutzenden Publikationen
- Verknüpfung mit Projekt / anderen FD

K1.8 Anerkennung der wissenschaftlichen Leistung

- Anreize für Forscherx
- Intregation Personal Identifier
- Zitationsraten / Altmetrics
- Reporting an Förderer

K2.2 Embargofunktion

- Embargo möglich
- Embargofrist
- Begründung & Ende sichtbar

K2.4 Integration in Zitationsdatenbanken

- Lieferung von MD an Zitationsdatenbanken
- Metadaten-Schnittstelle

- Art der API
- Funktion

K3 Transparenz und Qualitätssicherung

K3.1 Policies

- Richtlinien

K3.2 Datenvalidierung

- Vollständigkeit der MD
- Verantwortung für Inhalte

K3.3 Versionierung

- Versionierung möglich
- Anzeige u. Verknüpfung
- Änderungsprotokoll
- Persistent ID-Vergabe

K3.4 Löschung

- Löschung ist geregelt
- URL & PID bleiben aktiv
- Grund der Löschung angezeigt

K3.5 Sicherstellung langfristiger Zugänglichkeit

- Übernimmt Verantwortung für LZA
- Sicherungsmaßnahmen
- Sicherungsdauer
- Nachfolgeregelung bei Schließung

K3.6 Zertifizierung

- Ist zertifiziert
- Prüfberichte Online (nicht in Anhang [A.3](#) enthalten da keines der FDR zertifiziert)

K4 Nutzererlebnis und Support

K4.1 Usability

- Nutzerfreundlichkeit
- Intuitive Bedienung
- Etablierte Darstellungsformen
- Flache Lernkurve
- Barrierefreiheit

K4.2 Darstellung der Datensätze

- Datenvisualisierung
- Vorschaufunktion

K4.3 Support zu Repositoriumsaktionen

- Hilfeseiten, FAQs
- Schulungsangebote
- Supportanfragen
- Möglichkeit Fehler zu melden

K4.4 Beratung zum Forschungsdatenmanagement

- Beratung
- Hilfestellung
- Verweis auf Richtlinien / Standards

K5 Organisation und Infrastruktur

K5.1 Organisationsform

- Art der Einbindung in Organisation
- Langfristigkeit der Organisationsform
- Träger bekannt

- Gemeinnützig / Keine Gewinnerorientierung

- K5.2 Personalausstattung und -qualifizierung
- Anzahl Mitarbeiterx
 - Qualifizierung Mitarbeiterx
- K5.3 Forschungs- und Entwicklungstätigkeiten
- Weiterentwicklung der Infrastruktur
 - Initiativen des FDM und der Fachdisziplin
 - Arbeit an FDM-Standards
 - Forschungsprojekte
- K5.4 Technische Infrastruktur
- Allgemeine Informationen zur Software / Hardware
 - Lade- & Reaktionszeiten
 - Ausfallsicherheit durch verteilte Ressourcen
 - Back-Up Strategie
 - Überwachung der Systemkomponenten
 - technische Infrastruktur dokumentiert
 - Verwendung offener Software
- K6 Gebühren und Finanzierung
- K6.1 Gebühren für die Datenabgabe und -nutzung
- Gebührenmodell
 - Vertrag zwischen FDR und Institution / Träger
- K6.2 Finanzierung des Betriebs
- Langfristiges Finanzierungskonzept
 - Budgetplanung
 - Langfristige Finanzierungsprognose

Stanford Digital Repository

Beispieldatensätze für Screenshots:

Beispiel 1: <https://purl.stanford.edu/py542fc9766>

Beispiel 2: <https://purl.stanford.edu/qn021rf6349>

re3data Datensatz für das SDR [re317e].

K1.1 Datenabgabe

Zulässiges Format: Alle Formate sind zulässig [Sta Jd]. "We can take any type of data in any format." [Sta Jh]. Hat "Best Practice File Formats"²⁰² zusammengestellt.

Möglichkeiten des Uploads: SDR Online Deposit über Browser²⁰³, für größere Mengen Kontakt mit den Mitarbeiterx aufnehmen [Sta Jd].

Daten- Integritätsprüfung: Keine Hinweise auf Webseite gefunden.

K1.2 Metadaten

Metadatenkategorien: Werden nicht explizit getrennt behandelt.

standardisierte Formate: Beschreibende Metadaten in MODS und Dublin Core [Sta Jk], außerdem Verweis auf FGDC, EAD und DDI [Sta Ja]

unterstützende Tools: Zwei Webseiten: "Metadata Basics"²⁰⁴ und "Advanced Metadata"²⁰⁵.

²⁰²<https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-formats>

²⁰³<https://sdr.stanford.edu/>

²⁰⁴<https://library.stanford.edu/research/data-management-services/data-best-practices/creating-metadata/basic-approach-metadata>

²⁰⁵<https://library.stanford.edu/research/data-management-services/data-best-practices/creating-metadata/advanced-metadata>

Liste von Werkzeugen²⁰⁶ zusammengestellt und Vergleich in einer Tabelle²⁰⁷. Aktuell: "testing the use of CEDAR templates for creating RDF-encoded metadata describing digital artifacts" [Egy+17], mehr Info zu CEDAR Template Tools²⁰⁸

formularbasierte Eingabe: Ja, im Browser [Fro13]

Hinweis auf Pflichtangaben: Ja, required und optional [Fro13]

Hilfe bei Schlagworten (Systematik): Freie Vergabe von Schlagworten [Fro13, Minute 08:00].

Haben ein eigenes Metadata Department²⁰⁹

1.3 Dokumentation

Dokumentation der FD: Dokumentation kann als "related item" oder als eigene Datei gemeinsam mit Forschungsdaten abgelegt werden ([Fro13, Minute 9:20]).

Verlinkung auf Datenpublikation: Ist generell möglich über "related items", Bsp. für eine Verlinkung siehe Screenshot SDR002_RelatedItems.

K1.4 Zitierfähigkeit

Richtlinien zur Zitation von FD: Keine explizite Richtlinie, aber Zitationsvorschlag auf Datensatzebene (siehe Screenshot SDR001_PreferredCitation), vom Nutzerx generiert aus Metdaten oder selbst formatiert [Fro13]

PID: Nutzt PURL ("persistent URL") mit DRUID ("digital resource unique identifier") [Sta Jk]. "The PURL contains a persistent identifier called a DRUID (digital resource unique identifier) that functions much like a DOI." [Sta Jh]. "[...] an SDR PURL represents the crux of our digital library, providing a persistent, citable, online location for digital resources archived in SUL's preservation repository" [Fro15]

K1.5 Lizenzmodelle

Nutzungsbedingungen festlegen: Lizenz auf Datensatzebene textuell und visualisiert, siehe Screenshot SDR001_PreferredCitation

offene Lizenzen anbieten: Nutzt CC und ODC, "Creative Commons licenses, a matrix of six options, from liberal to conservative, for any kind of content. The SDR also offers Open Data Commons licenses." [Sta Jb]

Datenschutz: Keine Hinweise auf Webseite.

MD Bindungslos (CC0): Keine Hinweise auf Webseite.

K1.6 Verweise auf Veröffentlichungen

Link zu Veröffentlichung der Publikation: "data sets in SDR and link them directly with their publications" [Sta Jf]

Link zu Publikationen die nachnutzen: Nicht explizit erwähnt, aber möglich über Related Content Funktion.

Verknüpfung zugehöriger Projekte: Eingabe von Related Content im Rahmen der Einspeisung [Fro13, Minute 9:19].

K1.7 Nutzungsstatistik

Metriken auf Datensatzebene: Bisher nicht implementiert [Sta Je]

Standardisierte Erhebung: Nicht zutreffend.

K1.8 Anerkennung der wissenschaftlichen Leistung

Anreize für Forscherx: "Catalog record in SearchWorks, Stanford's primary access point for library and information resources." [Sta Jd], Featured Collections [Sta Jf], SDR Deposit of the

²⁰⁶<https://library.stanford.edu/research/data-management-services/data-best-practices/creating-metadata/metadata-tools>

²⁰⁷<https://library.stanford.edu/research/data-management-services/data-best-practices/creating-metadata/metadata-tools-comparison>

²⁰⁸<https://metadatacenter.org/tools-training/cedar-template-tools>

²⁰⁹<http://library.stanford.edu/departments/metadata-department>

Week auf dem Blog der Universitätsbibliothek²¹⁰.

Intregation PID: "No, the SDR does not currently support ORCID, but it is of increasing interest and appeal." [Sta Jk].

Zitationsraten / Altmetrics: Bisher nicht implementiert [Sta Je]

automatische Reportings: Keine Hinweise auf Webseite.

K2.1 Zugangsformen

öffentlich / Nutzergruppen: "[...] you can restrict access to a Stanford-only audience" [Sta Jb]. "offers two levels of content access: open to the public or restricted to Stanford affiliates (group-level restrictions is a future development)" [Sta JI]. Auf Sammlungsebene: Standard-einstellung ist "Everyone" also öffentlich, alternativ kann "Varies" eingestellt werden (Nutzer entscheidet Case-to-case ob Datensätze öffentlich sind) oder "Stanford community" (alle Datensätze einer Sammlung sind restricted) [Fro13].

Nutzerkreis Abgabe: "Members of the Stanford community – students, faculty, staff" [Sta Je]

Zugriffsteuerung (rechtliche / ethische Einschränkungen): "access is controlled using strict authentication policies and enterprise-level security mechanisms" [Sta Ji]

Metadaten restriktionsfrei: Keinen Hinweis oder Gegenbeweis auf Webseite gefunden.

K2.2 Embargofunktion

Embargofunktion: "control content release through embargo" [Sta Ji]. Embargo kann auf Sammlungsebene festgelegt werden (Standard ist "no delay" / kein Embargo, "Varies" erlaubt Sammlungsweite Maximaldauer festzulegen, "Fixed" setzt Embargo für alle Datensätze) [Fro13].

Embargofristen: Bis zu 3 Jahre [Sta Jd].

Begründung & Ende sichtbar: Keine Hinweise auf Webseite.

K2.3 Recherchemöglichkeiten

Allgemein: "The SDR does not have a stand-alone, dedicated search interface." stattdessen Einschränkung auf Kollektion über SearchWorks, dort seit 2013 eingebunden [Sta13].

Freie Textsuche: Ja, einfach und erweiterte Suche, siehe Screenshot SDR005_Suche.

Suchoperatoren: Boolesche (AND, OR, NOT), Trunkierung (*), Wildcard (?), Ausschliessen (-) und Einschliessen (+), siehe Screenshot SDR005_Suche.

attributspezifisch (Metadaten): Suche über Metadaten ist möglich, siehe Screenshot SDR006_SucheMD.

Facettierung: Über diverse Kategorien, siehe Screenshot SDR007_Facettierung.

Empfehlungssysteme: Keine Hinweise auf Webseite.

MD-Suche / Inhaltliche Suche: Nur MD, kein Hinweis auf inhaltliche Suche.

Indexierbare Landingpage: "Content is indexed into the library catalog, which is regularly crawled by Google, ensuring that your content will be discoverable via everyone's favorite search engine." [Hod17].

2.4 Integration in Zitationsdatenbanken

Metadaten bereit / liefert an Zit.-DB: Kein Hinweis auf Webseite.

MD-Abfrage-Schnittstellen (OAI-PMH): Kein Hinweis auf Webseite. Keine OAI-PMH Schnittstelle auf openarchives.org gelistet²¹¹.

K2.5 Programmierschnittstellen

APIs: Abfrage möglich, aber Ingest noch nicht: "A digital object with the appropriate visibility settings can be programmatically accessed via RESTful calls to its PURL (persistent URL) in

²¹⁰Beispiel: <https://library.stanford.edu/blogs/digital-library-blog/2017/05/sdr-deposit-week-video-tutorials-4d-visualization>

²¹¹<https://www.openarchives.org/Register/BrowseSites>

the SDR. The SDR does not yet have an exposed API for deposit [...]” [Sta Jk]. Hat seit 2016 eine IIIF Schnittstelle für Bilder [Sny16], und ist daran auch Mitentwickler²¹².

K3.1 Policies

Richtlinien: Nur "SDR Terms of Deposit"²¹³, sonst keine expliziten Richtlinien. Stattdessen verteilt auf den Hilfeseiten für Nutzerx, z.B. SDR Overview [Sta Ji], FAQ [Sta Je] usw.

K3.2 Datenvalidierung

Kein Hinweis auf Webseite.

K3.3 Versionierung

Versionierung: "Stanford Libraries developed Moab for versioning archived digital objects." [Sta Jj]. "Moab does not alter any content of object files [...] Manifests record file sizes and checksums for use in verification. [...] Any version of a digital object can be reconstructed and disseminated with equal ease to any other version." [And13, S. 29].

aktuelle V. gekennzeichnet / ältere V. verknüpft: "The most recent version will be available to users online; previous versions may be accessed upon request." [Sta Jd].

Änderungen protokolliert: "A version additions report lists which files were added to the object's catalog in a given version." [And13, S. 15].

PID-Vergabe deutlich: Nicht explizit auf Webseite.

Sonstiges: Video Tutorials "Create a new version of your published item in the Stanford Digital Repository"²¹⁴ und "Depositing to the Stanford Digital Repository"²¹⁵.

K3.4 Löschung

Geregelt: Keine explizite Policy. Löschen von einzelnen Dateien über Versionierung möglich [And13, S. 25].

URL & PID funktionsfähig: Kein Hinweis auf Webseite.

Grund der Löschung auf Landing-Page: Kein Hinweis auf Webseite.

K3.5 Sicherstellung langfristiger Zugänglichkeit

Übernimmt Verantwortung: "[...] enable Stanford researchers to preserve, manage, and share research data in a secure environment for long-term citation, access, and reuse." [Sta Jc].

Sicherungsmaßnahmen: "[...] highly-secure system where redundant copies are stored across storage systems in geodiverse locations. The SDR does not currently offer formal services in format migration or system emulation [...]" [Sta Jk]. "[...] the SDR currently manages over 230 TB of unique content (935,000 objects) using an enterprise tape library and ever-growing disk pool." [Sta Ji]. "[...] one preservation copy of SDR content is stored on spinning disk and multiple copies are stored on LTO tape. The tapes are stored in geo-diverse locations." [Sta Jj]. "Files are stored redundantly and are periodically audited to detect inadvertent change." [Sta Jd].

Sicherungsdauer: Kein Hinweis auf Webseite.

Nachfolgeregelungen: Keine explizite Nachfolgeregelung. "The SDR is a member of DPN, a federation of cultural heritage and higher education institutions working together to ensure a holistic and comprehensive approach to digital preservation." [Sta Jd].

K3.6 Zertifizierung

Nicht zertifiziert.

Einhaltung von Anforderungen: Nicht zutreffend.

Prüfberichte Online: Nicht zutreffend.

²¹² <https://library.stanford.edu/projects/international-image-interoperability-framework>

²¹³ <https://web.stanford.edu/group/sdr/SDR-Terms-of-Deposit-v13.pdf>

²¹⁴ <https://www.youtube.com/watch?v=CKXIAos65u0&t=11s>

²¹⁵ <https://youtu.be/F1T7guVkg8Q?t=14m50s>

K4.1 Usability

Nutzerxfreundlichkeit: Preferred Citation siehe Screenshot SDR001_PreferredCitation

intuitive Bedienung: Ja.

etablierte Darstellungsformen: Ja.

flache Lernkurve: Ja.

Barrierefreiheit: aXe zeigt 18 Verstöße auf Datensatzebene an (siehe Screenshot SDR008_Accessibility-Datensatz), vor allem Probleme mit Listenelementen in CSS (<dl>, <dt>), fehlende Beschriftung von Buttons, kein ausreichender Farbkontrast. Suche zeigt 50 Verstöße gegeben WCA Guidelines (siehe SDR009_Accessibility-Suche). Manuelle Betrachtung: Einige Videos haben Transkripte, aber nicht konsistent oder suchbar, siehe Screenshot (SDR010_Transkript). Keine Textalternativen (Stichprobenprüfung für 3 aktuelle Datensätze aus Resource Types Video, Image, Sound recording). Keine Anpassung Textgröße oder Farbschema. Grundsätzlich mit Tastatur allein bedienbar. Keine zeitlichen Einschränkungen. Keine Blitze. Keine Bemühungen hinsichtlich Lesbarkeit. Meist Vorhersehbar, aber Aufruf einer PURL zeigt andere Ansicht als Aufruf über SearchWorks. Keine Eingabehilfe oder Fehlerkorrektur.

K4.2 Darstellung der Datensätze

Datenvisualisierung: Keine Hinweise auf Datenvisualisierung.

Vorschaufunktion: Streaming auf der Webseite für Videos (siehe Screenshot SDR011_Vorschau) und Vorschau für Bilder.

K4.3 Support zu Repositoriumsaktionen

Hilfeseiten: Diverse Hilfeseiten²¹⁶, z.B. Guides und Videos²¹⁷.

FAQs: General²¹⁸, Copyright, licenses and access restrictions²¹⁹ und Technical topics²²⁰.

Schulungsangebote: Consulting, training, and other services²²¹.

Supportanfragen: Verlinkung auf Kontaktmöglichkeiten im Footer, siehe Screenshot SDR012_Supportanfrage.

Fehler melden: Nicht möglich.

K4.4 Beratung zum Forschungsdatenmanagement

Beratung: Data management services, Digitization services, Born-digital forensics services, General consultation services [Sta Jd]

Hilfestellung: Share and preserve research data²²², Managing Data FAQ²²³

Verweis auf aktuelle Richtlinien / Standards: Data best practices²²⁴, Liste disziplinspezifischer Repositorien²²⁵

K5.1 Organisationsform

²¹⁶<http://library.stanford.edu/research/stanford-digital-repository>

²¹⁷<http://library.stanford.edu/research/stanford-digital-repository/depositor-services/online-deposit>

²¹⁸<https://library.stanford.edu/research/stanford-digital-repository/faqs>

²¹⁹<http://library.stanford.edu/research/stanford-digital-repository/faqs/copyright-licenses-and-access-restrictions>

²²⁰<http://library.stanford.edu/research/stanford-digital-repository/faqs/technical-topics>

²²¹<https://library.stanford.edu/research/data-management-services/consulting-training-and-other-services>

²²²<https://library.stanford.edu/research/data-management-services/share-and-preserve/research-data>

²²³<https://library.stanford.edu/research/data-management-services/faq/managing-data-faq>

²²⁴<https://library.stanford.edu/research/data-management-services/data-best-practices>

²²⁵<https://library.stanford.edu/research/data-management-services/share-and-preserve/research-data/domain-specific-data-repositories>

Passt zu Zielen / Aufgabe: A bteilung der Stanford University Library [Sta Ji].

Langfristigkeit: Unbekannt, keine explizite Aussage auf Webseite.

Träger bekannt: Stanford University und Stanford University Library [Sta Ji].

Gemeinnützig: "unencumbered by commercial interests or proprietary systems" [Sta JI]

K5.2 Personalausstattung und -qualifizierung

"total staff of thirty – including service managers, preservation experts, software engineers, and system administrators" [Sta Ji].

K5.3 Forschungs- und Entwicklungstätigkeiten

Steuert Weiterentwicklung der Infrastruktur: "Stanford University partners with the Digital Public Library of America (DPLA) and DuraSpace on the Hydra-in-a-Box project, to develop, bundle, and promote a feature-rich, robust, flexible digital repository that is easier to install, configure, and maintain. This next-generation repository solution incorporates capabilities and affordances to support networked resources and services." [Sta17]

Kooperationen mit FDM- Initiativen und Fachdisziplin: "SUL also develops robust, custom-built websites for selected collections [...] SUL began building an application called Spotlight."

Plugin für Erstellung von Webseiten für digitale Sammlungen. [Sny14]

Mitarbeit an FDM-Standards: Manifesto for Linked Libraries²²⁶

Forschungsprojekte: Diverse Forschungsprojekte²²⁷.

K5.4 Technische Infrastruktur

Angemessen, Zuverlässig, Sicher, Skaliert: Hydra "[...] complete solution composed of open source technologies [...]", beinhaltet Fedora (Basis des Repositoriums), Solr (schnelle Suche), Blacklight "faceted searching, browsing and tailored views of objects" und Moab für Versionierung [Sta Jj].

schnelle Lade- & Reaktionszeiten: Ja.

Ausfallsicherheit (verteilte DB, Server): Keine weiteren Hinweise.

Back-Up Strategie: "In addition to the access copies readily available online, one preservation copy of SDR content is stored on spinning disk and multiple copies are stored on LTO tape. The tapes are stored in geo-diverse locations." [Sta Jj]

Überwachung der Komponenten: Keine weiteren Hinweise.

techn. Infrastrukt. dokumentiert: "For more information about the SDR's underlying technology, please contact the SDR team." [Sta Jj].

offene Software: Hydra "[...] complete solution composed of open source technologies [...]" [Sta Jj].

Forschungseth. und rechtl. Richtlinien: Keine expliziten Hinweise.

K6.1 Gebühren für die Datenabgabe und -nutzung

Bekannte Gebührenmodelle (Dauer / Volumen Staffellung): "currently in the process of establishing an overall pricing model for SDR Services" [Sta Jg].

Verträge zwischen FDR und Institut: Keine Hinweise auf Webseite.

K6.2 Finanzierung des Betriebs

Langfristiges Konzept: Keine Hinweise auf Webseite.

Budgetplanung: Keine Hinweise auf Webseite.

Prognosen für Zukunft: Keine Hinweise auf Webseite.

²²⁶ <https://library.stanford.edu/projects/linked-data>

²²⁷ <https://library.stanford.edu/projects>

Phaidra

Beispieldatensätze für Screenshots:

Beispiel 1: <https://hdl.handle.net/11353/10.448565>

Beispiel 2: <https://hdl.handle.net/11353/10.561325>

re3data Datensatz für Phaidra [re317d].

K1.1 Datenabgabe

Zulässiges Format: "alle Formate sind zulässig" (bevorzugt werden empfohlene Formate) [Abt Jh]. "Phaidra bietet Möglichkeit Texte (auch Bücher), Videos, Audios, Container, Collections und Sonderformate zu archivieren. Phaidra ist auf 'multi content resource type' vorbereitet (Modelle für multimediale Objekte, oder Primary Data, oder z.B. mathematische Formeln, etc.)" [Abt Jf]. "[...] digitale Speicherung, Verwaltung und Langzeitarchivierung von Medienobjekten, insbesondere von Grafik-, Text-, Video- und Audiodateien aus Forschung, Lehre und Wissenschaft." [Ami J].

Möglichkeiten des Uploads: "Single-Upload wird unterstützt, für Massen-Uploads sind sogenannte Docked Applications notwendig" [Gan17, S. 111]. "Wahl zwischen Einzelupload, Bulkupload in Zusammenarbeit mit dem ZID oder Nutzung des PhaidralImporters." [Blu16, S.4]. PhaidralImporter²²⁸

Daten-Integritätsprüfung: Keine Hinweise auf Webseite.

K1.2 Metadaten

Metadatenkategorien: "Der Metadateneditor ist in 9 Module aufgeteilt: - Allgemeine Angaben - Lebenszyklus des Objektes - Technische Angaben - eLearning Bereich (Educational) - Rechte und Lizenzen - Beziehung zwischen den Objekten - Anmerkungen - Klassifikation der Objekte - Zuordnungen zu den Organisationseinheiten." [Abt08, S. 24].

standardisierte Formate: Metadatenschema basiert auf Dublin Core und LOM [Abt Jg]. "UW-Metadaten, das auf dem LOM-Schema basiert; METS" [Gan17, S. 111]. "Als Metadaten-Format wird u.a. Dublin Core und das an die Universität Wien angepasste LOM-Schema verwendet." [Abt Jf]

unterstützende Tools: Docked Applications, "Applikation, die sich über die Schnittstelle mit Phaidra verbindet [...]. Diese Programme können für den Ingest (Daten nach Phaidra laden) [...] verwendet werden." [Gan17, S. 109].

formularbasierte Eingabe: "Metadateneingabe: manuell und auch über die API manipulierbar" [Gan17].

Hinweis auf Pflichtangaben: "In Phaidra sind derzeit folgende Metadatenfelder verpflichtend auszufüllen: Titel des Objekts, Beschreibung, Sprache des Objekts, Status, Rolle/n, Kosten, Copyright, Lizenz" [Abt Jc].

Hilfe bei Schlagworten (Systematik): "Dazu bietet die Software umfassende Werkzeuge für die formale und inhaltliche Erschließung, wie auch die Erstellung von E-Books (auch born digital). Die Verwendung von Schlagworten, Abstracts, Beschreibungen und Klassifikationen und Thesauri wie ÖFOS, PACS, EUROVOC; Getty; ACM; Basisklassifikation ist möglich." [Abt Jf]. Schlagworte sind unter "Klassifikation" angeordnet, aber nicht klickbar (siehe Screenshot PHA003_Sammlungen). In den erweiterten Metdaten wird die genutzte Klassifikation angezeigt (siehe Screenshot PHA005_Klassifikation). "Erstellung einer Vorlage (Template): Sie können Ihre eingegebenen Metadaten speichern und bei einer erneuten Eingabe wieder abrufen und verwenden." [Abt08, S. 25]

1.3 Dokumentation

Dokumentation der FD: Keine Hinweise zur Dokumentation von FD auf Webseite. Projekt

²²⁸ <https://phaidra.univie.ac.at/view/o:471960>

"code4research": "In Forschungsprojekten wird oft Software entwickelt, die nicht nur für die Datengenerierung entwickelt, sondern auch für die Darstellung der Daten eingesetzt wird. Im Fall der Darstellung und Repräsentation der Daten ist es notwendig, die Software auch nach Ende des Forschungsprojekts und somit auch nach Ende der Finanzierung weiter zu betreiben. [...] Unsere Idee ist es, ein Netzwerk an der Universität Wien aufzubauen um diese Fragestellungen möglichst umfassend zu diskutieren, um Wissen und Erfahrung auszutauschen und um ForscherInnen während der gesamten Projektlaufzeit Unterstützung anbieten zu können." [Abt Jd].

Verlinkung auf Datenpublikation: Keine Hinweise auf Webseite.

K1.4 Zitierfähigkeit

Richtlinien zur Zitation von FD: Keine Hinweise auf Webseite.

PID: "dauerhafte Adressen (z.B. handle, bei Büchern URN, für Publikationen und Forschungsdaten DOI)" [Abt Jg]. "Die digitalen Objekte werden mit individuellen und dauerhaften Adressen (Persistent Identifiers) versehen [...] Phaidra ermöglicht persistente Zitierbarkeit von Inhalten" [Abt Jf].

K1.5 Lizenzmodelle

Nutzungsbedingungen festlegen: Über die Terms of Use²²⁹ und Policy: "Das rechtliche Verhältnis zum Final-User wird durch die Verwendung von Lizenzen geregelt [...] Die Einhaltung von Urheber- und Verwertungsrechten Dritter liegt in der Verantwortung der LieferantInnen von Objekten." [Abt Jf].

offene Lizenzen anbieten: "[...] LieferantInnen von Objekten wird die Verwendung von Creative-Commons-Lizenzen nahe gelegt." [Abt Jf]. Liste von Lizenzen²³⁰

Datenschutz: "Der/die NutzerIn versichert, dass er/sie sich bei der Inanspruchnahme der Dienste des Diensteanbieters rechtmäßig verhält, insbesondere keine rechtswidrigen Zwecke verfolgt, keine Urheberrechtsverletzungen, keine Persönlichkeitsrechtverletzungen begeht und keine diskriminierenden, keine extremistischen und keine pornographischen Medienobjekte und Informationen verbreitet." [Ami J].

MD Bindungslos (CC0): "Die Metadaten sämtlicher Phaidra Objekte sind öffentlich verfügbar." [Abt Jg].

K1.6 Verweise auf Veröffentlichungen

Link zu Veröffentlichung der Publikation: "In einem weiteren Schritt entstehen evtl. Publikationen, die in u:scholar, dem Institutional Repository, einer Docked Application von Phaidra, hochgeladen werden können und in Phaidra langzeitarchiviert sind." [Abt Je].

Link zu Publikationen die nachnutzen: Keine Hinweise auf Webseite.

Verknüpfung zugehöriger Projekte: "Sämtliche Objekte können in Phaidra auf unterschiedliche Weise miteinander verknüpft werden." [Abt Jg].

K1.7 Nutzungsstatistik

Metriken auf Datensatzebene: "Der Diensteanbieter zeichnet bei jedem Seitenbesuch Webserver-Logdateien auf. Die Logdateien werden von der Systemadministration des Diensteanbieters insbesondere zur Fehleranalyse, bibliometrische Analyse, Download-Statistiken u.a. sowie zur Optimierung des Systems verwendet. Die Nutzung der Logdateien erfolgt ausschließlich nach deren Anonymisierung." [Ami J]. Anzahl der Aufrufe wird auf Datensatzebene angezeigt, siehe Screenshot PHA002_AnzeigeVideo.

Standardisierte Erhebung: Keine Hinweise auf Webseite.

K1.8 Anerkennung der wissenschaftlichen Leistung

Anreize für Forscherx: Keine Hinweise auf Webseite.

²²⁹ https://phaidra.univie.ac.at/terms_of_use/show_terms_of_use

²³⁰ <https://datamanagement.univie.ac.at/rdm/juristisches/lizenzmodelle/>

Intregation PID: Keine Hinweise auf Webseite.

Zitationsraten / Altmetrics: Anzahl der Aufrufe wird auf Datensatzebene angezeigt, siehe Screenshot PHA002_AnzeigeVideo.

automatische Reportings: Keine Hinweise auf Webseite.

K2.1 Zugangsformen

öffentlich / Nutzergruppen: "Offener und freier Zugang (alle MitarbeiterInnen und alle Studierende der Universität Wien können Phaidra aktiv nutzen) [...] Differenziertes Zugriffskonzept" [Abt Jg]. "Es wird zwischen aktiver (z.B. Upload von Objekten) und passiver Verwendung von Phaidra (Ansehen und Download der Objekte) unterschieden. [...] Grundsätzlich sind alle Objekte, die in Phaidra gespeichert werden, weltweit ohne Registrierung sichtbar und nutzbar. Durch das differenzierte Zugriffskonzept von Phaidra kann es jedoch sein, dass Digitalisate gesperrt sind." [Abt Ji]. "Die Nutzerkreise können über Gruppen definiert werden, und es gibt vordefinierte Gruppen, die sich aus der Zugehörigkeit zu Organisationseinheiten ergeben." [Gan17, S. 109].

Nutzerkreis Abgabe: "Aktive Nutzergruppen: MitarbeiterInnen (per u:account), Studierende (per u:account), Für Projekte fordern Sie bitte eine Servicemailadresse an, u:accounts für Externe (bitte 'erweiterte Rolle' wählen)" [Abt Ji]. "[...] außeruniversitären Instituten steht Phaidra Depot zur Verfügung." [Abt Je] (in dieser Arbeit nicht betrachtet).

Zugriffsteuerung (rechtliche / ethische Einschränkungen): "Die Nutzerkreise können über Gruppen definiert werden, und es gibt vordefinierte Gruppen, die sich aus der Zugehörigkeit zu Organisationseinheiten ergeben." [Gan17, S. 109]. Technisch über u:account²³¹ System der Universität Wien.

Metadaten restriktionsfrei: "Die Metadaten sämtlicher Phaidra Objekte sind öffentlich verfügbar." [Abt Jg].

K2.2 Embargofunktion

Embargofunktion: "[...] eigene Objekte für einen eingeschränkten Nutzerkreis zugänglich zu machen oder generell zu sperren." [Abt Jf]. "Läuft die Sperrfrist ab, so wird das Objekt automatisch öffentlich zur Verfügung gestellt. Die Metadaten der Objekte sind immer öffentlich zugänglich." [Gan17, S. 109].

Embargofristen: "Embargozeiten und Sperrfristen können vom Datenlieferant festgelegt werden." [Abt Jf].

Begründung & Ende sichtbar: Kein Hinweis auf Webseite gefunden.

K2.3 Recherchemöglichkeiten

Allgemein: Freie Textsuche: "Die einfache Suche geht über alle Metadatenfelder, die über die Expertensuche eingeschränkt werden kann." [Gan17, S. 111].

Suchoperatoren: Keine Suchhilfe auf Webseite. Stichprobentests für Wildcards u. Trunkierung (siehe Screenshot PHA010_Suchfunktionen-Trunkierung-Wildcard) und Boolesche Operatoren (siehe Screenshot PHA011_Suchfunktionen-Bool) scheinen funktionsfähig.

attributspezifisch (Metadaten): "Phaidra bietet strukturierte Metadaten und stellt deshalb auch metadatenbasierte Suchstrategien zur Verfügung." [Abt Jf].

Facettierung: "Es gibt auch eine BlätternFunktion auf Objekttypen." [Gan17, S. 111]. Facettierung siehe Screenshot PHA001_Typen von Daten.

Empfehlungssysteme: Kein Hinweis auf Webseite gefunden.

MD-Suche / Inhaltliche Suche: Keine Hinweise auf Inhaltliche / Volltextsuche.

Indexierbare Landingpage: "Die auf Phaidra gespeicherten Medienobjekte sind insbesondere mittels ihrer Metadaten durch Suchmaschinen, etwa durch 'google' auffindbar." [Ami J].

²³¹<http://zid.univie.ac.at/uaccount-benutzungsordnung/>

2.4 Integration in Zitationsdatenbanken

Metadaten bereit / liefert an Zit.-DB: Kein Hinweis auf Webseite gefunden.

MD-Abfrage-Schnittstellen (OAI-PMH): "Metadaten über OAI-PMH, Daten können auch über die API exportiert werden." [Gan17, S. 112]. "Die weitere Übertragung von Daten erfolgt über einen OAI-MPH Provider." [Abt Jf]. Base-URL wird nicht auf der Webseite angezeigt.

K2.5 Programmierschnittstellen

APIs: "REST (JSON), OAI-PMH, LTI in Planung / Metadaten über OAI-PMH, Daten können auch über die API exportiert werden." [Gan17, S. 111]. Dokumentation der phaidra-api²³² auf github. "Docked Application ist ein Begriff von Phaidra und bezeichnet eine Applikation, die sich über die Schnittstelle mit Phaidra verbindet und Objekte, die in Phaidra abgelegt sind, darstellen. Diese Programme können für den Ingest (Daten nach Phaidra laden) und auch für den Re-Use (Nachnutzung, die Daten in einer speziellen Form darstellen) verwendet werden." [Gan17, S. 109]. "[...] Vielfalt standardisierter Schnittstellen zur flexiblen Weiterverwendung der verwalteten Ressourcen und Objekte." [Abt Jf].

K3.1 Policies

Richtlinien: Eine explizite Policy²³³, zusätzlich Terms of Use²³⁴. Universität Wien hat außerdem eine OA Policy²³⁵.

K3.2 Datenvalidierung

Metadatenfeld für Angabe zu durchgeführtem Peer Review wird unter erweiterten Metadaten angezeigt (siehe Screenshot PHA006_PeerReview), keine weiteren Hinweise.

K3.3 Versionierung

Versionierung: "Objekte können nicht verändert werden, es ist jedoch möglich neue Versionen zu erstellen." [Abt Jf].

aktuelle V. gekennzeichnet / ältere V. verknüpft: "Die Vorgängerversion bleibt mit ihrem Persistent Identifier erhalten und wird an das neue Objekt verlinkt." [Abt Jf]. Version wird auf Datensatzebene angezeigt, siehe Screenshot PHA002_AnzeigeVideo.

Änderungen protokolliert: Keine Hinweise auf Webseite.

PID-Vergabe deutlich: "Die Vorgängerversion bleibt mit ihrem Persistent Identifier erhalten und wird an das neue Objekt verlinkt." [Abt Jf]

K3.4 Löschung
Geregelt: "Einmal in Phaidra hochgeladen, ist ihr Objekt permanent gespeichert, d.h. Sie können es nicht löschen." [Abt08, S. 19]. "Lädt ein/eine Nutzerin ein Medienobjekt hoch, kann er/sie grundsätzlich dessen Löschung durch den Diensteanbieter im Nachhinein nicht verlangen. Es sei denn er/sie macht einen schwer wiegenden Grund geltend, dem nicht durch eine Sperrung des Objekts abgeholfen werden kann." [Ami J].

URL & PID funktionsfähig: Keine Hinweise auf Webseite.

Grund der Löschung auf Landing-Page: Keine Hinweise auf Webseite.

K3.5 Sicherstellung langfristiger Zugänglichkeit

Übernimmt Verantwortung: "Der Diensteanbieter beabsichtigt Datenbestände zu erhalten und der Flüchtigkeit von Daten entgegen zu wirken." [Ami J].

Sicherungsmaßnahmen: "Für die Erschließung, Speicherung und Archivierung der Objekte werden internationale Standards wie die Richtlinien der Open Archives Initiative genutzt und weiterentwickelt." [Abt Jf].

"Ewige Migration ist zur Zeit der Ansatz für die langfristige Zurverfügungstellung der Daten.

²³² <https://github.com/phaidra/phaidra-api/wiki/Documentation>

²³³ <http://datamanagement.univie.ac.at/ueber-phaidra/policy-von-phaidra/>

²³⁴ https://phaidra.univie.ac.at/terms_of_use/show_terms_of_use

²³⁵ <https://openaccess.univie.ac.at/open-access/oa-policy-der-uni-wien/>

Zu berücksichtigen sind alle drei Ebenen." (Daten: Bitstreampreservation / Physisches Medium, Formate: Überführung in aktuelle, Semantik und Logik: Prüfung nach jeder Migration) [BG17].

Sicherungsdauer: "[...] Dauer von mindestens fünf Jahren ab dem Zeitpunkt der Speicherung." [Ami J].

Nachfolgeregelungen: "Für den Fall einer aus wichtigem, unabwendbarem Grund erfolgenden Einstellung des Dienstes verpflichtet sich der Diensteanbieter die an ihn übermittelten Medienobjekte im Rahmen seiner Möglichkeiten zu erhalten und auf Verlangen eine digitale Kopie dessen dem/der NutzerIn, der/die sie hochgeladen hat, zurück zu übermitteln." [Ami J].

K3.6 Zertifizierung

Nicht zertifiziert.

Einhaltung von Anforderungen: Nicht zutreffend.

Prüfberichte Online: Nicht zutreffend.

K4.1 Usability

Nutzerfreundlichkeit: "Erfassung der Metadaten nicht sehr benutzerfreundlich, und das Suchinterface ist nicht klar." [Gan17, S. 111]. "Die Verarbeitung sämtlicher Sprachen ist gewährleistet. Das User-Interface ist mehrsprachig (stand 2012): Deutsch, Englisch, Italienisch, Serbisch." [Abt Jf].

intuitive Bedienung: Ja.

etablierte Darstellungsformen: Ja.

flache Lernkurve: Ja.

Barrierefreiheit: Browsererweiterung aXe findet 11 Verstöße gegen Barrierefreiheit (siehe Screenshot PHA012_Accessibility-Suche). 8 Verstöße auf Datensatzebene (siehe Screenshot PHA013_Accessibility-Datensatz), bei beiden vor allem fehlender Farbkontrast, aber auch fehlende Beschriftung (Buttons, Form-tags, IDs). Manuelle Betrachtung: Die Beschreibung von Videos in den Metadaten kann per Freitext umgesetzt werden, siehe PHA004_Accessibility, dies ist aber nur optional / nicht normiert. Keine Funktion für das Einstellen von Transkript oder Untertitel gefunden. Keine Textalternativen (Stichprobenprüfung für jeweils 3 Datensätze aus Typ Video, Bild, Audio). Keine Anpassung Textgröße oder Farbschema. Nicht mit Tastatur allein bedienbar. Keine zeitlichen Einschränkungen. Keine Blitze. Keine Bemühungen hinsichtlich Lesbarkeit. Eingabehilfe siehe Screenshot PHA014_Accessibility-Suchvorschläge. Keine Fehlerkorrektur.

K4.2 Darstellung der Datensätze

Datenvisualisierung: "In Phaidra gibt es Plug-ins, über die spezielle Inhaltstypen von Phaidra aus besser dargestellt werden können. So wird zum Beispiel der Typ „Book“ mit dem „Book viewer“ dargestellt." [Gan17, S. 110]. "Für die Visualisierung der Daten stehen mehrere Möglichkeiten zur Verfügung. Das Content Management System UNIDAM ermöglicht weitere Präsentationsmöglichkeiten." [Abt Jb].

Vorschaufunktion: Streamingoption für Videos, Ansicht auf Webseite statt Download möglich (siehe Screenshot PHA002_AnzeigeVideo).

K4.3 Support zu Repositoriumsfunktionen

Hilfeseiten: Keine explizite Hilfeseite oder allgemeine FAQs, aber Informationsseiten zu Themen wie Technik²³⁶ oder Zugang²³⁷.

Schulungsangebote: Übersicht aktuelle Schulungen und Workshops²³⁸. "Trainings und Work-

²³⁶ <http://datamanagement.univie.ac.at/ueber-phaidra/technik/>

²³⁷ <http://datamanagement.univie.ac.at/ueber-phaidra/zugang-zu-phaidra/>

²³⁸ <http://datamanagement.univie.ac.at/>

shops werden durch die Bibliothek angeboten oder werden im Rahmen der universitären Fortbildungsmaßnahmen an der Personalentwicklung durchgeführt. Zusätzlich werden bei Bedarf auch individuelle Termine angeboten.” [Abt Jf].

Supportanfragen: ”Unterstützt werden die NutzerInnen durch ein Customer Management (erreichbar unter phaidra@univie.ac.at). Bei technischen Problemen stehen der Phaidra-Service zur Verfügung, der unter support.phaidra@univie.ac.at erreichbar ist und der universitätsweite Help Desk des ZIDs der Universität Wien.” [Abt Jf].

Fehler melden: Kein Hinweis auf Webseite.

K4.4 Beratung zum Forschungsdatenmanagement

Beratung: ”Die Abteilung Phaidra Unidam unterstützt Sie bei Fragen zu der Eingabe/automatischen Übernahme von Metadaten in Phaidra und Unidam, bei Mappingfragen, beim Ausfüllen eines Datenmanagementplans und beim Vernetzen mit Expertinnen.” [Abt Jc].

Hilfestellung: ”Hilfe bei der Projektantragsstellung, Projektbegleitung, Anpassung der Infrastruktur nach Möglichkeiten, Metadaten, Juristisches, Unterstützung beim Ingest.” [BG17, S. 23]. ”[...] allen Usern mit u:account das Testsystem, die Phaidra-Sandkiste zur Verfügung.” [Abt Je]. Guideline für Datenmanagementpläne²³⁹. Haftungsszenarien in einem Online-Repository²⁴⁰ zusammengestellt. Verweis auf aktuelle Richtlinien / Standards: Nutzt das FAIR Prinzip [BG17].

K5.1 Organisationsform

Passt zu Zielen / Aufgabe: Abteilung Phaidra - UNIDAM²⁴¹ ist Teil der Universitätsbibliothek. Langfristigkeit: Kein Hinweis auf Webseite.

Träger bekannt: ”Phaidra wurde an der Universität Wien entwickelt und wird vom Zentralen Informatikdienst (ZID) technisch und von der Universitätsbibliothek inhaltlich betreut. Die Projektleitung erfolgt an der Universitätsbibliothek.” [Abt Jf].

Gemeinnützig: Non-Profit laut re3data [re317d].

K5.2 Personalausstattung und -qualifizierung

Nur 3 Mitarbeiterx genannt auf Webseite²⁴². Keine Hinweise zur Qualifizierung.

K5.3 Forschungs- und Entwicklungstätigkeiten

Steuert Weiterentwicklung der Infrastruktur: ”Der Betrieb und die Weiterentwicklung der auf Fedora Commons Repository basierenden Software erfolgt an der Universität Wien.” [Abt Jf].

”OpenAIRE ist ein paneuropäisches Forschungsinformationssystem mit Diensten zum Auffinden, Speichern, Verlinken und Analysieren von Forschungsergebnissen über alle Disziplinen hinweg. [...] Gesamtziel des Projekts [e-Infrastructures Austria] ist der koordinierte Aufbau und die Weiterentwicklung von Repositorieninfrastrukturen in ganz Österreich.” [Abt Ja].

Kooperationen mit FDM- Initiativen und Fachdisziplin: ”Das Netzwerk für RepositorienmanagerInnen (Repmanet) ist ein Zusammenschluss von Personen aus Institutionen in ganz Österreich, die ein Repository betreuen und sich mit anderen über Erfahrungen, Herausforderungen und Spezialthemen austauschen möchten. [...] Größen-, inhalts- und technikenabhängig soll ein fachlicher Austausch über unterschiedliche Themen stattfinden und nach möglichen Strategien und Lösungen für offene Fragen gesucht werden.” [BF J].

Mitarbeit an FDM-Standards: Keine Hinweise auf Webseite.

Forschungsprojekte: ”Open Education Austria ist ein Projekt österreichischer Universitäten zur gemeinsamen Entwicklung einer nationalen Infrastruktur für Open Educational Resour-

²³⁹ <https://datamanagement.univie.ac.at/rdm/datenmanagementplaene/>

²⁴⁰ <https://datamanagement.univie.ac.at/rdm/juristisches/faqhaftungsszenarien/>

²⁴¹ <http://datamanagement.univie.ac.at/>

²⁴² <http://datamanagement.univie.ac.at/ueber-uns/>

ces (OER). [...] Der Kulturpool bietet einen zentralen Zugang zu digitalen österreichischen Kulturerbe-Ressourcen und verweist auch auf Objekte, die in Phaidra archiviert sind." [Abt Ja]. Diverse Publikationen²⁴³.

K5.4 Technische Infrastruktur

Angemessen, Zuverlässig, Sicher, Skaliert: "Suche basierend auf apache solr" [Abt Jg]. "Phaidra läuft in einer Linux Umgebung und braucht einen Apache- und einen Tomcat-Server. Der Fedora-Teil ist in JAVA entwickelt und der Phaidra-Core in Perl. Teile des Frontends sind in AngularJS entwickelt." [Gan17, S. 110]. "Phaidra ist ein digitales Archiv und am OAIS Referenzmodell orientiertes Repository zur Verwaltung und Bereitstellung von digitalen Ressourcen auf Basis von FEDORA." [Abt Jf]. Zusammenfassung der technischen Basis auf Webseite²⁴⁴.

schnelle Lade- & Reaktionszeiten: Keine Auffälligkeiten, keine explizite Aussage auf Webseite.
Ausfallsicherheit (verteilte DB, Server): Keine Hinweise auf Webseite.

Back-Up Strategie: Keine Hinweise auf Webseite.

Überwachung der Komponenten: Keine Hinweise auf Webseite.

techn. Infrastrukt. dokumentiert: Dokumentation Online in github²⁴⁵, auch für API²⁴⁶.

offene Software: Alle verwendeten sind typischerweise offene Software, aber kein Hinweis auf Webseite.

K6.1 Gebühren für die Datenabgabe und -nutzung

Bekannte Gebührenmodelle: "Offener und freier Zugang" [Abt Jg]. "Die elektronische Veröffentlichung und Archivierung ist für die LieferantInnen von Objekten kostenfrei [...]" [Abt Jf]. "Der Diensteanbieter stellt unentgeltlich Kapazität zum Speichern von Medienobjekten auf einem mit dem Internet verbundenen Server zur Verfügung." [Ami J].

Verträge zwischen FDR und Institut: Keine Hinweise auf Webseite.

K6.2 Finanzierung des Betriebs

Langfristiges Konzept: Keine Hinweise auf Webseite oder in Jahresberichten der Universitätsbibliothek²⁴⁷.

Budgetplanung: Keine Hinweise gefunden.

Prognosen für Zukunft: Keine Hinweise gefunden.

Zenodo

Beispieldatensätze für Screenshots:

Beispiel 1: <https://zenodo.org/record/1099523>

Beispiel 2: <https://zenodo.org/record/1079032>

re3data Datensatz für Zenodo [re317f].

K1.1 Datenabgabe

Zulässiges Format: "All formats are allowed - even preservation unfriendly. [...] Total files size limit per record is 50GB." [Zen Jd].

Möglichkeiten des Uploads: Formular, siehe Screenshot Zen008_Upload und Schnittstelle: "Zenodo REST API currently supports: Deposit — upload and publishing of research outputs (identical to functionality available in the user interface)." [Zen Jh].

Daten-Integritätsprüfung: "Data quality: All information is provided "as-is", and the user shall

²⁴³<http://datamanagement.univie.ac.at/ueber-phaidra/publikationen/>

²⁴⁴<http://datamanagement.univie.ac.at/ueber-phaidra/technik/>

²⁴⁵<https://github.com/phaidra>

²⁴⁶<https://github.com/phaidra/phaidra-api/wiki/Documentation>

²⁴⁷http://bibliothek.univie.ac.at/zahlen_und_daten.html

hold Zenodo and information providers supplying data to Zenodo free and harmless in connection with the use of such information." [Zen Jd].

K1.2 Metadaten

Metadatenkategorien: Nicht explizit getrennt.

standardisierte Formate: "All metadata is stored internally in JSON-format according to a defined JSON schema. Metadata is exported in several standard formats such as MARCXML, Dublin Core, and DataCite Metadata Schema (according to the OpenAIRE Guidelines)." [Zen Jd]. "Metadata is stored in JSON format in PostgreSQL in a structure described by versioned JSONSchemas. All changes to metadata records on Zenodo are versioned, and happening inside database transactions." [Zen Je]. "Zenodo's metadata is compliant with DataCite's Metadata Schema minimum and recommended terms, with a few additional enrichments." [Zen Jg].

unterstützende Tools: Keine Hinweise auf Webseite.

formularbasierte Eingabe: Ja, siehe Screenshot Zen008_Upload.

Hinweis auf Pflichtangaben: "Each record contains a minimum of DataCite's mandatory terms, with optionally additional DataCite recommended terms and Zenodo's enrichments." [Zen Jg].

Hilfe bei Schlagworten (Systematik): Freitext, keine Systematik hinterlegt für Schlagworte (siehe Screenshot Zen009_Schlagworte). Aber andere Felder: "For certain terms we refer to open, external vocabularies, e.g.: license (Open Definition), funders (FundRef) and grants (OpenAIRE)." [Zen Jg].

1.3 Dokumentation

Dokumentation der FD: "Zenodo makes the sharing, curation and publication of data and software a reality for all researchers. [...] Data, software and other artefacts in support of publications may be the core [...]" [Zen Ja]. "The addition [of versioning] is part of our efforts to make software a citable, first-class citizen in the world of scientific research." [loa17].

Verlinkung auf Datenpublikation: "Specify identifiers of related publications and datasets." [Zen Jf] im Reiter "Related/alternate identifiers" (siehe Screenshot Zen003_VerwandtePublikationen).

K1.4 Zitierfähigkeit

Richtlinien zur Zitation von FD: Zitationsvorschlag auf Datensatzebene (siehe Screenshot Zen011_Zitationsvorschlag). Unterstützt Data Citation Principles²⁴⁸.

PID: "A DOI is issued to every published record on Zenodo. [...] The DOI is a top-level and a mandatory field in the metadata of each record." [Zen Jg]. Für die Verlinkung zu anderen Publikationen: "Supported identifiers include: DOI, Handle, ARK, PURL, ISSN, ISBN, PubMed ID, PubMed Central ID, ADS Bibliographic Code, arXiv, Life Science Identifiers (LSID), EAN-13, ISTC, URNs and URLs." [Zen Jf].

K1.5 Lizenzmodelle

Nutzungsbedingungen festlegen: "Users must specify a license for all publicly available files. Licenses for closed access files may be specified in the description field." [Zen Jd]. "Zenodo encourages you to share your research as openly as possible to maximize use and re-use of your research results. [...] Therefore, we allow for uploading under a variety of different licenses and access levels." [Zen Jb].

offene Lizenzen anbieten: Nutzt Creative Commons (siehe Screenshot Zen012_Lizenzen).

Datenschutz: "The uploader shall ensure that their content is suitable for open dissemination, and that it complies with these terms and applicable laws, including, but not limited to, privacy, data protection and intellectual property rights." [Zen Ji].

MD Bindungslos (CC0): "Metadata are publicly accessible and licensed under public domain. No authorization is ever necessary to retrieve it." [Zen Jg]. "All meta data is openly available

²⁴⁸<https://www.force11.org/datacitationprinciples>

under CC0 licence [...]” [Zen Ja].

K1.6 Verweise auf Veröffentlichungen

Link zu Veröffentlichung der Publikation: Möglich über "Specify identifiers of related publications and datasets." [Zen Jf] im Reiter "Related/alternate identifiers" (siehe Screenshot Zen003_VerwandtePublikationen).

Link zu Publikationen die nachnutzen: Möglich über "Specify identifiers of related publications and datasets." [Zen Jf] im Reiter "Related/alternate identifiers" (siehe Screenshot Zen003_VerwandtePublikationen).

Verknüpfung zugehöriger Projekte: "Zenodo allows you to create your own collection [...]" [Zen Jb]. Verknüpfung über Communities, siehe Screenshot Zen002_Communities.

K1.7 Nutzungsstatistik

Metriken auf Datensatzebene: Nur Altmetrics, siehe K 1.8.

Standardisierte Erhebung: Keine Hinweise auf Webseite.

K1.8 Anerkennung der wissenschaftlichen Leistung

Anreize für Forscherx: "Zenodo is a strong supporter of open data in all its forms (meaning data that anyone is free to use, reuse, and redistribute) and takes an incentives approach to encourage depositing under an open license. We therefore only display Open Access uploads on the front-page." [Zen Jc].

Intregation PID: Direkte Verknüpfung mit Orcid Konto möglich (siehe Screenshot Zen013_Orcid). "Last but not least, you can now include an author's ORCID under the Authors section of the metadata on the deposit web interface." [Now17].

Zitationsraten / Altmetrics: "Article Level Metric" Beispiel auf Poster [NS14]. Siehe Screenshot Zen004_AltMetric, von Beispieldatensatz: <https://doi.org/10.5281/zenodo.837902>. Achtung! Bei Bearbeitung keine Anzeige in Chrome Version 63.0.3239.84 (Stand 29.12.17), Altmetrics nur im Inkognito Modus sichtbar.

automatische Reportings: "Zenodo helps researchers receive credit by making the research results citable and through OpenAIRE integrates them into existing reporting lines to funding agencies like the European Commission. Citation information is also passed to DataCite and onto the scholarly aggregators." [Zen Ja].

K2.1 Zugangsformen

öffentlich / Nutzergruppen: "Files may be deposited under closed, open, or embargoed access. Files deposited under closed access are protected against unauthorized access at all levels. [...] Users may deposit restricted files with the ability to share access with others if certain requirements are met. These files will not be made publicly available and sharing will be made possible only by the approval of depositor of the original file." [Zen Jd]. "Access to Zenodo's content is open to all, for non-military purposes only." [Zen Ji]. "Zenodo allows users to upload files under closed access. Closed access means that zenodo.org users will not be able to access the files you uploaded. The files are however stored unencrypted and may be viewed by Zenodo operational staff under specific conditions. This means that "closed access" on Zenodo is not suitable for secret or confidential data." [Zen Je].

Nutzerkreis Abgabe: "Anyone may register as user of Zenodo. All users are allowed to deposit content [...]" [Zen Jd]. "Zenodo is part of OpenAIRE and provides a repository for those researchers who do not have an existing institutional or thematic repository they can deposit their publications and data in." [Zen Jc].

Zugriffsteuerung (rechtliche / ethische Einschränkungen): "Zenodo allows users to upload files under closed access. Closed access means that zenodo.org users will not be able to access the files you uploaded. The files are however stored unencrypted and may be viewed by Zenodo

operational staff under specific conditions. This means that “closed access” on Zenodo is not suitable for secret or confidential data.” [Zen Je].

Metadaten restriktionsfrei: “Metadata are publicly accessible and licensed under public domain.” [Zen Jg]. “All meta data is openly available under CC0 licence [...]” [Zen Ja].

K2.2 Embargofunktion

Embargofunktion: “Users may deposit content under an embargo status and provide and [sic!] end date for the embargo. The repository will restrict access to the data until the end of the embargo period; at which time, the content will become publically [sic!] available automatically.” [Zen Jd]. “Content can also be embargoed and automatically opened when the associated paper is published.” [Zen Ja].

Embargofristen: Kein Hinweis auf Webseite.

Begründung & Ende sichtbar: Keine Begründung, aber Ende des Embargo auf Datensatzebene sichtbar (siehe Screenshot Zen014_Embargo).

K2.3 Recherchemöglichkeiten

Freie Textsuche: Einfache Schlitzsuche (siehe Screenshot Zen005_Suchschlitz).

Suchoperatoren: Keine Hilfeseite, keine Aussagekräftigen Ergebnisse bei Tests mit Trunkierung, Wildcards, BOOL möglich.

attributspezifisch (Metadaten): “Metadata is in addition indexed in an Elasticsearch cluster for fast and powerful searching.” [Zen Je]. “Metadata of each record is indexed and searchable directly in Zenodo’s search engine immediately after publishing.” [Zen Jg].

Facettierung: Facettierung (siehe Screenshot Zen006_Facettierung) Access Right, File Type, Keywords, Type.

Empfehlungssysteme: Kein Hinweis auf Webseite.

MD-Suche / Inhaltliche Suche: Nur MD-Suche, kein Hinweis auf inhaltliche Suche.

Indexierbare Landingpage: Keine Hinweise auf Webseite, aber eine Known-Item-Search in Google ist erfolgreich (siehe Screenshot Zen016_IndexingGoogle).

2.4 Integration in Zitationsdatenbanken

Metadaten bereit / liefert an Zit.-DB: “Metadata of each record is sent to DataCite servers during DOI registration and indexed there.” [Zen Jg]. “Citation information is also passed to DataCite and onto the scholarly aggregators.” [Zen Ja].

MD-Abfrage-Schnittstellen (OAI-PMH): “Metadata for individual records as well as record collections are harvestable using the OAI-PMH protocol by the record identifier and the collection name. Metadata is also retrievable through the public REST API.” [Zen Jg]. “All metadata is exported via OAI-PMH and can be harvested.” [Zen Jd].

K2.5 Programmierschnittstellen

APIs: “Access to metadata and data files is provided over standard protocols such as HTTP and OAI-PMH.” [Zen Jd]. “Metadata for individual records as well as record collections are harvestable using the OAI-PMH protocol by the record identifier and the collection name. Metadata is also retrievable through the public REST API.” [Zen Jg]. “The Zenodo REST API currently supports: Deposit — upload and publishing of research outputs [...] The following REST APIs are currently in testing before we launch them in beta with full documentation: Records — search published records. Files — download/upload of files. Communities - search communities. Funders — search for funders. Grants — search for grants. Licenses — search for licenses.” [Zen Jh].

K3.1 Policies

Richtlinien: “Zenodo does not sign SLAs (service-level agreements). Instead, Zenodo is run by leading practitioners according to best practices.” Nutzt FAIR Principles [Zen Jg]. Hat Terms

of Use²⁴⁹ und Policies²⁵⁰.

K3.2 Datenvalidierung

"Data quality: All information is provided "as-is", and the user shall hold Zenodo and information providers supplying data to Zenodo free and harmless in connection with the use of such information." [Zen Jd]. "The uploader shall ensure that their content is suitable for open dissemination, and that it complies with these terms and applicable laws, including, but not limited to, privacy, data protection and intellectual property rights." [Zen Ji].

K3.3 Versionierung

Versionierung: "we have extended our metadata schema with the version term, so you can explicitly specify the version of your software, dataset, or any other type of upload." [loa17]. "Data files are versioned. Records are not versioned. The uploaded data is archived as a Submission Information Package. Derivatives of data files are generated, but original content is never modified. Records can be retracted from public view; however, the data files and record are preserved." [Zen Jd].

aktuelle V. gekennzeichnet / ältere V. verknüpft: "When you publish an upload on Zenodo for the first time, we register two DOIs: - a DOI representing the specific version of your record. -a DOI representing all of the versions of your record. Afterwards, we register a DOI for every new version of your upload." [Zen Jc].

Änderungen protokolliert: Keine Hinweise auf Webseite.

PID-Vergabe deutlich: Eigenes Kapitel "DOI Versioning" in den FAQs²⁵¹.

K3.4 Löschung

Geregelt: "Withdrawal is considered an exceptional action, which normally should be requested and fully justified by the original uploader." [Zen Jd].

URL & PID funktionsfähig: "The DOI and the URL of the original object are retained." [Zen Jd].

Grund der Löschung auf Landing-Page: "If the uploaded research object must later be withdrawn, the reason for the withdrawal will be indicated on a tombstone page, which will henceforth be served in its place." [Zen Jd].

K3.5 Sicherstellung langfristiger Zugänglichkeit

Übernimmt Verantwortung: Siehe Kapitel longevity in den General Policies²⁵².

Sicherungsmaßnahmen: "All data files are stored in CERN Data Centres, primarily Geneva, with replicas in Budapest. [...] multiple replicas in a distributed file system, which is backed up to tape on a nightly basis. [...] no promises of usability and understandability of deposited objects over time [...] stored along with a MD5 checksum of the file content. Files are regularly checked against their checksums." [Zen Jd]. "All files uploaded to Zenodo are stored in CERN's EOS service in an 18 petabytes disk cluster. Each file copy has two replicas located on different disk servers. [...] Metadata and persistent identifiers in Zenodo are stored in a PostgreSQL instance operated on CERN's Database on Demand infrastructure with 12-hourly backup cycle with one backup sent to tape storage once a week. All changes to metadata records on Zenodo are versioned, and happening inside database transactions. [...] Zenodo may, depending on access patterns in the future, move the archival and/or the online copy to CERN's offline long-term tape storage system CASTOR in order to minimize long-term storage costs." [Zen Je].

Sicherungsdauer: "Items will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined

²⁴⁹<http://about.zenodo.org/terms/>

²⁵⁰<http://about.zenodo.org/policies/>

²⁵¹<http://help.zenodo.org/>

²⁵²<http://about.zenodo.org/policies/>

for the next 20 years at least." [Zen Jd].

Nachfolgeregelungen: "In the highly unlikely event that Zenodo will have to close operations, we guarantee that we will migrate all content to other suitable repositories, and since all uploads have DOIs, all citations and links to Zenodo resources (such as your data) will not be affected." [Zen Jc]. "In case of closure of the repository, best efforts will be made to integrate all content into suitable alternative institutional and/or subject based repositories." [Zen Jd].

K3.6 Zertifizierung

Nicht zertifiziert.

Einhaltung von Anforderungen: Nicht zutreffend.

Prüfberichte Online: Nicht zutreffend.

K4.1 Usability

Nutzerxfreundlichkeit: Defekte Fehlermeldung: Zen015_Usability-Error (Eingabe in der Suche mit Anführungszeichen).

intuitive Bedienung: Ja.

etablierte Darstellungsformen: Ja.

flache Lernkurve: Beeinträchtigt, keine Suchhilfe.

Barrierefreiheit: Browsererweiterung aXe findet 173 Verstöße gegen Barrierefreiheit in Anzeige der Suchergebnisse (siehe Screenshot Zen017_Accessibility-Suche), vor allem fehlender Farbkontrast und fehlende Beschriftung in form-tags. 46 Verstöße auf Datensatzebene (siehe Screenshot Zen018_Accessibility-Datensatz), vor allem fehlender Farbkontrast und fehlende Textalternativen und Linkbeschriftungen. Manuelle Betrachtung: Beschreibung von Videos ist optional, nur über Freitext vorgesehen. Keine Funktion für das Einstellen von Transkript oder Untertitel gefunden. Keine Textalternativen (Stichprobenprüfung für jeweils 3 Datensätze aus Type Image (Photo), Video/Audio). Keine Anpassung Textgröße oder Farbschema. Nicht mit Tastatur allein bedienbar. Keine zeitlichen Einschränkungen. Keine Blitze. Keine Bemühungen hinsichtlich Lesbarkeit. Keine Eingabehilfe und keine Fehlerkorrektur in der Suche.

K4.2 Darstellung der Datensätze

Datenvisualisierung: Keine Hinweise auf Webseite.

Vorschaufunktion: Keine Vorschaufunktion u. kein Streaming von Videos. Vorschaufunktion für einzelne Bilder möglich (siehe Screenshot Zen019_PreviewBild).

K4.3 Support zu Repositoriumsaktionen

Hilfeseiten: Frequently Asked Questions²⁵³.

Schulungsangebote: Kein Hinweis auf Webseite.

Supportanfragen: Über das Contact Form²⁵⁴.

Fehler melden: Kein Hinweis auf Webseite.

K4.4 Beratung zum Forschungsdatenmanagement

Beratung: Kein Hinweis auf Webseite.

Hilfestellung: Kein Hinweis auf Webseite.

Verweis auf aktuelle Richtlinien / Standards: Kein Hinweis auf Webseite.

K5.1 Organisationsform

Passt zu Zielen / Aufgabe: "Zenodo is embedded in the IT Department, Collaboration Devices and Applications Group, Digital Repositories Section (IT-CDA-DR)." [Zen Je]. Webseite der Abteilung²⁵⁵.

²⁵³ <http://help.zenodo.org/>

²⁵⁴ <https://zenodo.org/support>

²⁵⁵ <http://information-technology.web.cern.ch/about/organisation/digital-repositories>

Langfristigkeit: "Zenodo is hosted by CERN which has existed since 1954 and currently has an experimental programme defined for the next 20+ years." [Zen Je].

Träger bekannt: "Zenodo is funded by: European Commission via the OpenAIRE projects: FP7: OpenAIRE (246686), OpenAIREplus (283595), Horizon 2020: OpenAIRE2020 (643410) and OpenAIRE-Connect (731011), CERN, Alfred P. Sloan Foundation, Donations via CERN & Society Foundation." [Zen Je].

Gemeinnützig: Keine Hinweise auf Webseite.

K5.2 Personalausstattung und -qualifizierung

"CERN has some of the world's top experts in running large scale research data infrastructures and digital repositories that we rely on in order to deliver a trusted digital repository. [...] Zenodo is however embedded in a much larger team, headed by Jose Benito Gonzalez Lopez [...]" [Zen Je]. Liste der aktuellen Mitarbeiter (15) in der zuständigen Abteilung²⁵⁶ "CERN has considerable knowledge and experience in building and operating large scale digital repositories" [Zen Jc].

K5.3 Forschungs- und Entwicklungstätigkeiten

Steuert Weiterentwicklung der Infrastruktur: "We are partners in multiple European Commission funded projects, amongst others: OpenAIRE, EUDAT." [Zen Je].

Kooperationen mit FDM- Initiativen und Fachdisziplin: Keine Hinweise auf Webseite.

Mitarbeit an FDM-Standards: Joint Declaration of Data Citation Principles²⁵⁷.

Forschungsprojekte: "CERN is an active member of the following organisations and international bodies (non-exhaustive): DataCite, ORCID, FORCE11 (in particular Data Citation Principles and Software Citation Principles), Research Data Alliance, SCOAP3." [Zen Je].

K5.4 Technische Infrastruktur

Angemessen, Zuverlässig, Sicher, Skaliert: Aktueller Status von Zenodo²⁵⁸ (Files upload/download, Search, Website, Uptime, Quickstats). Aktueller Stand 30.12.2017 siehe Screenshot Zen020_Status. "Zenodo frontend servers are responsible for running the Invenio repository platform application which is based on Python and the Flask web development framework. The frontend servers are running nginx HTTP server and uwsgi application server in front of the application and nginx is in addition in charge of serving static content." [Zen Je]. "Your research output is stored safely for the future in same cloud infrastructure as research data from CERN's Large Hadron Collider and using CERN's battle-tested repository software INVENIO, which is used by some of the world's largest repositories such as INSPIRE HEP and CERN Document Server." [Zen Jb].

schnelle Lade- & Reaktionszeiten: Update vom 12.09.17: "Accuracy and speed have been significantly improved thanks to the brand new search backend powered by Elasticsearch, coupled with a new search user interface." [Zen17] und Screenshot Zen021_Sept17Update (da Webseite nach jedem Update garantiert neu).

Ausfallsicherheit (verteilte DB, Server): "We are furthermore running two independent systems: one production system and one quality assurance system. This ensures that all changes, whether at infrastructure level or source code level, can be tested and validated on our quality assurance system prior to being applied to our production system." [Zen Je].

Back-Up Strategie: "Data files and metadata are backed up nightly and replicated into multiple copies in the online system [...] Metadata and persistent identifiers in Zenodo are stored in a PostgreSQL instance operated on CERN's Database on Demand infrastructure with 12-hourly backup cycle with one backup sent to tape storage once a week." [Zen Je].

²⁵⁶ <https://phonebook.cern.ch/phonebook/#search/?query=IT-CDA-DR>

²⁵⁷ <https://www.force11.org/datacitation/endorsements>

²⁵⁸ <https://status.zenodo.org/>

Überwachung der Komponenten: "Servers are monitored via CERN's monitoring infrastructure based on Flume, Elasticsearch, Kibana and Hadoop. Application errors are logged and aggregated in a local Sentry instance. Traffic to Zenodo frontend servers is load balanced via a combination of DNS load balancing and HAProxy load balancers." [Zen Je].

techn. Infrastrukt. dokumentiert: APIs dokumentiert²⁵⁹ und Zenodo SourceCode auf github²⁶⁰.

offene Software: "Zenodo is powered by CERN Data Centre and the Invenio digital library framework and is fully run on open source products all the way through." [Zen Je].

Forschungseth. und rechtl. Richtlinien: Kein Hinweis auf Webseite.

K6.1 Gebühren für die Datenabgabe und -nutzung

Bekannte Gebührenmodelle (Dauer / Volumen Staffelung): "Content may be uploaded free of charge by those without ready access to an organized data centre." [Zen Ji]. Im Kontext von Datensätzen über 50GB: "Please be aware that we cannot offer infinite space for free, so donations towards sustainability are encouraged." [Zen Jc].

Verträge zwischen FDR und Institut: Keine Hinweise auf Webseite.

K6.2 Finanzierung des Betriebs

Langfristiges Konzept: "Items will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years at least." [Zen Jd].

Budgetplanung: Keine Hinweise auf Webseite.

Prognosen für Zukunft: Keine Hinweise auf Webseite.

Allgemein: Giving to CERN (Spendenseite) beinhaltet auch Zenodo²⁶¹.

Brown Digital Repository

Beispieldatensätze für Screenshots:

Beispiel 1: <https://doi.org/10.7301/Z0668B3H>

Beispiel 2: <https://doi.org/10.7301/Z0TB14TR> (Link zum Video), <https://repository.library.brown.edu/studio/item/bdr:432042/> (Link zum Datensatz).

re3data Datensatz für das BDR [re316a].

K1.1 Datenabgabe

Allgemein: "[...] offers support for storing, retrieving, and curating: Publications, Research data, Digital objects that inform scholarly work [...] Individuals have a 4GB limit when uploading personal collections. Users requiring more space should contact the Library [...] Examples of appropriate materials include: Publications, Data sets, Audio, video, or images derived from research or creative works, Research documentation, Computer programs developed through research, Art work, Digital surrogates of primary sources." [Bro Jf].

Zulässiges Format: "There are no restrictions on the file types that you can upload. Most common types of files (e.g., PDF, DOC, XML, JPG, etc.) will be recognized correctly by the system. File types that are not recognized will be stored and downloaded intact, but will show an "unknown file type" icon." [Bro Je]. "The following file formats will be recognized by the BDR and made accessible via the web in their native format: pdf, MSWord (doc, docx, etc.), ppt, Excel (xls, xlsx, etc.), txt, jp2, jpg, jpeg, tiff, zip, jar, gz, flv, mov, mp4, mp3, xml, xslt" [Bro Jf].

²⁵⁹ <http://developers.zenodo.org>

²⁶⁰ <https://github.com/zenodo/zenodo>

²⁶¹ <https://giving.web.cern.ch/civicrm/contribute/transact?reset=1&id=20>

Möglichkeiten des Uploads: "Upload Screen" wird in den FAQ erwähnt [Bro Je].

Daten-Integritätsprüfung: "Objects added to the digital repository will be subject to review by Library staff to ensure data integrity, metadata standards compliance, and system security." [Bro Jf].

K1.2 Metadaten

Metadatenkategorien: Auf den Hilfeseiten der Bibliotheksabteilung "Digital Production Services" wird Descriptive, Administrative, Technical, Structural, Preservation und Rights metadata unterschieden [Bro Ji].

standardisierte Formate: "The BDR uses a combination of widely-implemented metadata standards, including: - Dublin Core [...] - MODS [...] - METS [...] - Other common metadata standards may also be used." [Bro Jf].

unterstützende Tools: keine Hinweise auf Webseite.

formularbasierte Eingabe: "Upload Screen" wird in den FAQ erwähnt [Bro Je]. Keine weiteren Hinweise.

Hinweis auf Pflichtangaben: Keine Hinweise auf Webseite.

Hilfe bei Schlagworten (Systematik): Keine Systematik, aber mehr freie Schlagwörter: "Additionally, user defined tags can be added through a web-based tool." [Bro Jf].

1.3 Dokumentation

Dokumentation der FD: Forschungsdokumentation explizit einbezogen: "Examples of appropriate materials include: Publications, Data sets, Audio, video, or images derived from research or creative works, Research documentation, Computer programs developed through research, Art work, Digital surrogates of primary sources." [Bro Jf].

Verlinkung auf Datenpublikation: Möglich über Metadatum "Related Item" siehe Screenshot BDR004_Verlinkung oder unter Full Metadata²⁶².

K1.4 Zitierfähigkeit

Richtlinien zur Zitation von FD: Zitationsvorschlag auf Datensatzebene (siehe Screenshot BDR003_Zitationsvorschlag).

PID: "Each object that is submitted to the BDR will be made accessible to the submitter via a permanent link." [Bro Jf]. "The Brown Digital Repository is pleased to offer the ability to mint Digital Object Identifiers (DOI) for members of the Brown Community. [...] All students, faculty, and staff may request a DOI for digital assets produced via the scholarly, instructional, research, and administrative activities at Brown." [Bro Jd].

K1.5 Lizenzmodelle

Nutzungsbedingungen festlegen: "Authors/creators retain copyright. Authors/creators grant the Library non-exclusive license to preserve all items, and to distribute items based on their access levels. Author may still reuse content elsewhere." [Bro Jf].

offene Lizenzen anbieten: CC und GNU (siehe Screenshots BDR010_Lizenz-CC und BDR011_Lizenz-GNU).

Datenschutz: Keine Hinweise auf Webseite.

MD Bindungslos (CC0): Keine Hinweise auf Webseite.

K1.6 Verweise auf Veröffentlichungen

Link zu Veröffentlichung der Publikation: Keine Hinweise auf Webseite, grundsätzlich möglich über Metadatum "Related Item" siehe Screenshot BDR004_Verlinkung.

Link zu Publikationen die nachnutzen: Keine Hinweise auf Webseite, grundsätzlich möglich über Metadatum "Related Item" siehe Screenshot BDR004_Verlinkung.

Verknüpfung zugehöriger Projekte: Sammlungen siehe Screenshot BDR006_Sammlungen.

K1.7 Nutzungsstatistik

²⁶² Beispieldatensatz: <https://doi.org/10.7301/Z0668B3H>

Metriken auf Datensatzebene: Entwicklung eines Pilotprojekts für Metriken: "Collect and report metrics on holdings and usage of the BDR" [Bro Jj, S. 3].

Standardisierte Erhebung: Nicht zutreffend.

K1.8 Anerkennung der wissenschaftlichen Leistung

Anreize für Forscher: Nennt Vorteile des DOI Service als "Benefits for Your Research Product", "DOIs allow you the ability to cite your own research products in your publications, so that others may better discover them and then provide you with proper attribution in their scholarship." [Bro Jd].

Intregation PID: Keine Hinweise auf Nutzung im BDR, aber "Brown is a member of ORCID which allows the University, among other things, to create ORCID records on behalf of faculty, students, and affiliated individuals [...]" [Rai15].

Zitationsraten / Altmetrics: Über den Umweg DOI-Vergabe: [...] allowing you to use Web tools such as ImpactStory, Google Scholar, Google Analytics, etc. to measure the impact of your research products published on the Web. You can use these tools to help you collect analytics to visualize how and how often your digital research products are being discovered and used by others." [Bro Jd].

automatische Reportings: Keine Hinweise auf Webseite.

K2.1 Zugangsformen

öffentlich / Nutzergruppen: "The submitter has the ability to control basic access levels (Public, Brown University only, Private) through web-based tools. More flexible permissions (e.g., groups of people, departments, collaborators, etc.) can be set up as needed." [Bro Jf]. "In the "Set Privacy" options in the Upload Screen you may choose, Private (only I can see this), Brown (Everyone at Brown can see this), or Public (Everyone can see this). These settings can be changed by you at any time and will be instantly updated." [Bro Je].

Nutzerkreis Abgabe: "As of May 2012, the BDR is open to submissions by all faculty. Student work may be included in the BDR as part of faculty or administrative projects." [Bro Jf].

Zugriffsteuerung (rechtliche / ethische Einschränkungen): More flexible permissions (e.g., groups of people, departments, collaborators, etc.) can be set up as needed. [...] Access is controlled using the University's Shibboleth authentication system." [Bro Jf].

Metadaten restriktionsfrei: Keine Hinweise auf Webseite.

K2.2 Embargofunktion

Embargofunktion: "We have been storing dissertations in the BDR for a while. Students have the option to embargo their dissertations, and in that case we set the access rights so that the dissertation documents are only accessible to the Brown community (although the metadata is still accessible to everyone)." [Cai17b]. Keine Aussage zu FD in BDR.

Embargofristen: Nicht zutreffend.

Begründung & Ende sichtbar: Nicht zutreffend.

K2.3 Recherchemöglichkeiten

Allgemein: Keine Sortierung der Treffer (über Zeit oder Autor o.ä.) siehe Screenshot BDR007_Suche. Außerdem über Schnittstelle: Search API nutzt Solr query [Bro Jc].

Freie Textsuche: Einfache Suche über alle Felder oder Title, Author, Subject (siehe Screenshot BDR007_Suche).

Suchoperatoren: Keine Suchhilfe. Test mit BOOL, Wildcard und Trunkierung ohne Fehlermeldung, Suchergebnisse nicht beweiskräftig (siehe Screenshot BDR012_WildcardTrunkierung). Test mit exakten Suchbegriffen erfolgreich (siehe Screenshot BDR013_ExactSearch).

attributspezifisch (Metadaten): Mindestens Title, Author, Subject (siehe Screenshot BDR007_Suche).

Facettierung: Über Format, Keywords und Genre. Genre ist potentiell Freitext, siehe Screen-

shot BDR008_Facetten.

Empfehlungssysteme: Keine Hinweise auf Webseite.

MD-Suche / Inhaltliche Suche: Entwicklung eines Pilotprojekts für Volltextsuche und "Expand support for research data collections (enable larger more diverse datasets) in the BDR." [Bro Jj, S. 1].

Indexierbare Landingpage: "Will my uploaded files show up in Google? If you make them publicly available in the Privacy settings they will, otherwise no." [Bro Je]. "All metadata may be shared with search engines, other repositories, or through RSS or Atom feeds." [Bro Jf].

2.4 Integration in Zitationsdatenbanken

Metadaten bereit / liefert an Zit.-DB: "All metadata may be shared with search engines, other repositories, or through RSS or Atom feeds." [Bro Jf]. Keine weiteren Hinweise auf Webseite.

MD-Abfrage-Schnittstellen (OAI-PMH): Keine Hinweise auf Webseite.

K2.5 Programmierschnittstellen

APIs: Item API "Returns data about the object, such as primary_title and object_type." [Bro Jc], Collection API für Sammlungen und Search API "The search query should be formatted the same way as standard Solr queries." [Bro Jc].

K3.1 Policies

Richtlinien: Explizite Policy²⁶³. Außerdem über Brown University: "Copyright and Fair Use at Brown"²⁶⁴ und "Acceptable Use Policy"²⁶⁵ der IT-Abteilung.

K3.2 Datenvalidierung

"Objects added to the digital repository will be subject to review by Library staff to ensure data integrity, metadata standards compliance, and system security." [Bro Jf].

K3.3 Versionierung

Versionierung: "Changes to bitstreams will create new archived versions and an audit trail to monitor changes over time." [Bro Jf].

aktuelle V. gekennzeichnet / ältere V. verknüpft: Kein Hinweis auf Webseite.

Änderungen protokolliert: "Changes to bitstreams will create new archived versions and an audit trail to monitor changes over time." [Bro Jf].

PID-Vergabe deutlich: "New versions of existing objects will use the same permanent links unless an entirely new object is created." [Bro Jf].

K3.4 Löschung

Geregelt: "When individual [sic!] leaves Brown, the Library will remove items upon request unless they fall under existing policies that grant the University ongoing copyright or are required for compliance purposes." [Bro Jf]. Keine weiteren Hinweise auf Webseite.

URL & PID funktionsfähig: Kein Hinweis auf Webseite.

Grund der Löschung auf Landing-Page: Kein Hinweis auf Webseite.

K3.5 Sicherstellung langfristiger Zugänglichkeit

Übernimmt Verantwortung: "The Brown Digital Repository (BDR) is intended to make accessible and preserve digital objects of enduring value to the scholarly mission of Brown University." [Bro Jf]. "Permanent, secure storage for personal and departmental digital objects." [Bro Jb].

Sicherungsmaßnahmen: "Off-site backups of digital content. [...] Data curation, format migration, and preservation services." [Bro Jb]. "The bitstreams that make up an object will be maintained intact, with mechanisms available for checking their integrity. [...] Over time, some file formats may need to be migrated to more modern formats to remain viable, while others

²⁶³ <https://repository.library.brown.edu/studio/policies>

²⁶⁴ <https://www.brown.edu/about/administration/copyright>

²⁶⁵ <https://it.brown.edu/computing-policies/acceptable-use-policy>

may have no modern equivalent. [...] In the event that Library staff need to modify, migrate, or otherwise change the state of deposited items, the Library will notify the depositor prior to any changes. Every effort will be made to retain pre-migration data unless the depositor chooses to withdraw the items from the repository." [Bro Jf].

Sicherungsdauer: "Permanent, secure storage for personal and departmental digital objects." [Bro Jb]. Keine weiteren Hinweise auf Webseite.

Nachfolgeregelungen: Kein Hinweis auf Webseite.

K3.6 Zertifizierung

Nicht zertifiziert.

Einhaltung von Anforderungen: Nicht zutreffend.

Prüfberichte Online: Nicht zutreffend.

K4.1 Usability

Nutzerfreundlichkeit: Keine Suchhilfe.

intuitive Bedienung: Ja.

etablierte Darstellungsformen: Ja.

flache Lernkurve: Nein, keine Suchhilfe.

Barrierefreiheit: Browsererweiterung aXe findet 44 Verstöße gegen Barrierefreiheit in Anzeige der Suchergebnisse (siehe Screenshot BDR014_Accessibility-Suche), vor allem fehlender Farbkontrast und fehlende Beschriftung von Links. 3 Verstöße auf Datensatzebene (siehe Screenshot BDR015_Accessibility-Datensatz), vor allem fehlender Farbkontrast. Manuelle Betrachtung: Videos sind nicht inhaltlich beschrieben. Universität hat einen Closed Captioning Service ²⁶⁶ "Any video that will be widely distributed or featured on Brown's website should be captioned, as should course videos when requested by students or faculty members." [Bro Ja]. Keine Textalternativen (Stichprobenprüfung für jeweils 3 Datensätze aus Format Still image, Moving image und Sound recording). Keine Anpassung Textgröße oder Farbschema. Mit Tastatur allein bedienbar. Keine zeitlichen Einschränkungen. Keine Blitze. Keine Bemühungen hinsichtlich Lesbarkeit. Keine Eingabehilfe und keine Fehlerkorrektur in der Suche. In Planung für 2017/18: "- Support use of VPAT (Voluntary Product Accessibility Template) with library vendors and campus-wide, -Produce an internal values statement on accessibility. Add accessibility checks to all projects." [Bro Jj].

K4.2 Darstellung der Datensätze

Datenvisualisierung: Diverse Anzeigemodi für Bilder (siehe Screenshot BDR016_Viewer). Über CDS Abteilung der Bibliothek: "CDS has in house expertise using a number of data visualization tools and methods. [...] CDS is available to assist faculty, graduate students, and staff with projects that use Geographic Information Systems (GIS) for spatial analysis, data management, and producing static (printed) or dynamic (interactive web-based) maps." [Bro Jg].

Vorschaufunktion: Vorschaufunktion für Videos (siehe Screenshot BDR005_Vorschau) und Bilder (siehe Screenshot BDR016_Viewer). "You can upload any media file under the size limit and download it again. If you have media that you need to stream from the repository, please contact bdr@brown.edu for assistance." [Bro Je]. "We currently use OpenSeadragon in the BDR, and it's IIIF-compliant, but we've also looked at Universal Viewer and Mirador." [Cai16a].

K4.3 Support zu Repositoriumsaktionen

Hilfeseiten: Frequently Asked Questions ²⁶⁷, Informationen zu DOI ²⁶⁸ und APIs ²⁶⁹.

²⁶⁶ <https://it.brown.edu/services/type/closed-captioning-digital-content>

²⁶⁷ <https://repository.library.brown.edu/studio/faq>

²⁶⁸ <https://repository.library.brown.edu/studio/doi/>

²⁶⁹ <https://repository.library.brown.edu/studio/api-docs/>

Schulungsangebote: Workshops zu div. Themen, z.B. Datenvisualisierung ²⁷⁰

. Supportanfragen: Verlinkung "Contact the Digital Repository" in Fußleiste verweist auf e-Mail Adresse.

Fehler melden: Kein Hinweis auf Webseite.

K4.4 Beratung zum Forschungsdatenmanagement

Beratung: Support über das Center for Digital Scholarship: Data Curation, Visualization, Metadata Creation, Digital Humanities allgemein, Digitization, Workshops [Bro Jg]. Support über die Digital Production Services: Digitalisierung, inkl. Metadata consulting, Metadata application, Training [Bro Ji].

Hilfestellung: Über die Bibliothek: "The Library can help you to deposit data in the BDR [...]" [Bro Jk] und diverse "Data Management and Sharing Services", darunter z.B. Vorlagen für Datenmanagementpläne²⁷¹. Die Computing & Information Services (Bibliotheks-extern, Brown IT Abteilung) bieten u.a. Video Recording, Editing, and Delivery²⁷².

Verweis auf aktuelle Richtlinien / Standards: Verweise auf "Resources for Data Management Best Practices"²⁷³, Data Curation Best Practices²⁷⁴ und re3data²⁷⁵.

K5.1 Organisationsform

Passt zu Zielen / Aufgabe: Ja, Ziel ist "[...] place to gather, index, store, preserve, and make available digital assets produced via the scholarly, instructional, research, and administrative activities at Brown." [Bro Jb], FDR ist direkt an der Bibliothek angesiedelt.

Langfristigkeit: Keine Hinweise auf Webseite.

Träger bekannt: "The Brown University Library maintains the repository as a service to the Brown community [...]" [Bro Jb]. "The Library's Digital Technologies department develops and manages technological systems [...] The department is comprised of three areas: Integrated Technology Services, Brown Digital Repository services, Library Web Services." [Bro JI].

Gemeinnützig: Keine expliziten Hinweise auf Webseite.

K5.2 Personalausstattung und -qualifizierung

Keine Informationen zum BDR selbst auf Personalseite der Bibliothek ²⁷⁶. Die Abteilung Center for Digital Scholarship hat 8 Mitarbeiter²⁷⁷ inkl. einem Digital Repository Manager und einem Data Visualization Coordinator. Die Abteilung Digital Technologies hat 3 Mitarbeiter für das BDR²⁷⁸: Je ein Digital Repository Manager, Digital Library Programmer, Digital Preservation Librarian.

K5.3 Forschungs- und Entwicklungstätigkeiten

Steuert Weiterentwicklung der Infrastruktur: Über CDS: "Digital repositories and data curation: At the center of CDS's work is the development of durable repositories and best practices for long-term curation of scholarly research data." [Bro Jh].

Kooperationen mit FDM- Initiativen und Fachdisziplin: Diverse Digitalisierungsprojekte mit der Fachdisziplin²⁷⁹.

Mitarbeit an FDM-Standards: "CDS staff also serve in standards organizations and contribute

²⁷⁰ <https://library.brown.edu/cds/activities/workshops/#data-analysis>

²⁷¹ https://library.brown.edu/info/data_management_plans

²⁷² <https://it.brown.edu/services/type/video-recording-editing-and-delivery>

²⁷³ https://library.brown.edu/info/data_management_resources

²⁷⁴ https://library.brown.edu/info/data_management_curation

²⁷⁵ https://library.brown.edu/info/data_management_depositing

²⁷⁶ <https://library.brown.edu/about/deptlist.php#dept12>

²⁷⁷ <https://library.brown.edu/cds/about/people/>

²⁷⁸ <https://library.brown.edu/dt/>

²⁷⁹ <https://library.brown.edu/cds/projects/>

to the development of new standards. [...] Metadata standards and best practices: CDS takes an active role in the research community in this area, with CDS staff serving on standards development groups and providing testing for new tools and data formats." [Bro Jh].

Forschungsprojekte: " CDS projects include experimental digital monographs, digital journals, digital research collections, and hybrid publications that interrogate the boundaries between these forms. [...] Scholarly technology architecture and tools: We examine how emerging technologies enable and enhance digital research methodologies, and we build frameworks and end-user tools to advance digital scholarship efforts at Brown and elsewhere." [Bro Jh].

K5.4 Technische Infrastruktur

Allgemein: "Loris is a IIF-compliant image server, and its development is led by Jon Stroop at Princeton. Loris is written in Python, and it uses the WSGI standard." [Cai16b]. "We are currently using Fedora 3 for storing our repository object binaries and metadata. However, Fedora 3 is end of life and unsupported, so eventually we'll have to decide what our plan for the future is." [Cai17a]. Die Abteilung die das BDR technisch betreut hat "Digital Technologies Strategic Goals, 2017-18", z.B.: Entwicklung eines Pilotprojekts für Volltextsuche und "Expand support for research data collections (enable larger more diverse datasets) in the BDR." [Bro Jj, S. 1], sowie "Collect and report metrics on holdings and usage of the BDR" [Bro Jj, S. 3].

Angemessen, Zuverlässig, Sicher, Skaliert: Keine Hinweise auf Webseite.

schnelle Lade- & Reaktionszeiten: Keine Hinweise auf Webseite.

Ausfallsicherheit (verteilte DB, Server): "One big change was migrating from RHEL 5 to RHEL 7, but we also moved from basically a one-server setup to four separate servers. [...] We installed Loris on a new server [...] We also added a new server for Solr, and then a new server for Fedora. These two services previously ran on the one server that handled almost everything for the BDR [...] One advantage of being on four servers is the security we get from having our services isolated." [Cai17c].

Back-Up Strategie: Keine Hinweise auf Webseite.

Überwachung der Komponenten: Keine Hinweise auf Webseite.

techn. Infrastruktur. dokumentiert: Keine Hinweise auf Webseite.

offene Software: Linux (Red Hat Enterprise Linux, RHEL), Fedora, IIF, Loris sind offen. Sonst keine expliziten Hinweise auf Webseite.

Forschungseth. und rechtl. Richtlinien: Keine Hinweise auf Webseite.

K6.1 Gebühren für die Datenabgabe und -nutzung

Bekannte Gebührenmodelle (Dauer / Volumen Staffellung): Keine Hinweise auf Webseite.

Verträge zwischen FDR und Institut: Keine Hinweise auf Webseite.

K6.2 Finanzierung des Betriebs

Langfristiges Konzept: "Brown University Libraries Technology Plan, 2015-2017"²⁸⁰ enthielt Pläne für das BDR, bisher kein Hinweis auf Plan ab 2018.

Budgetplanung: Keine Hinweise auf Webseite.

Prognosen für Zukunft: Keine Hinweise auf Webseite.

LSE Research Online

Beispieldatensatz für Screenshots:

Beispiel: <http://eprints.lse.ac.uk/id/eprint/62869>

²⁸⁰ <https://library.brown.edu/dt/LORITechnologyPlan2015.pdf>

Beispiel: <http://eprints.lse.ac.uk/id/eprint/67380>, nur Altmetrics re3data Datensatz für LSERO [re317b].

K1.1 Datenabgabe

Allgemein: "We can accept research from LSE Academic and non-Academic staff members, as well as LSE PhD research students." [LSE Je]. "Visitors to LSERO benefit from free access to world class LSE research, including journal articles and book chapters which may otherwise be behind subscription paywalls. Equally discoverable are conference videos, presentations, technical reports and other 'grey literature' and nonpublished research outputs, which can often be difficult to locate online." [LSE14].

Zulässiges Format: Keine Hinweise auf Webseite. Bibliothek hat eine Liste mit empfohlenen offenen Formaten²⁸¹, darunter WebM für audiovisuelle FD [LSE Jb]. Typen von Ressourcen die angenommen werden: Articles – accepted / submitted / published version, Working Papers, Datasets, Books, Book Chapters, Conference items, Multimedia, Posters, Newspaper Articles Website items, Research Blogs [LSE Je].

Möglichkeiten des Uploads: "Email your research output or publication details to lseresearchonline@lse.ac.uk [...] Academic staff are welcome to add their own papers." [LSE Je].

Daten- Integritätsprüfung: "We rely on academics to provide details of publications they wish to include in the repository [...] Library staff will complete bibliographic details and add the full text." [LSE Je].

K1.2 Metadaten

Allgemein: Export in diverse Formate möglich, z.B. Dublin Core, METS (siehe Screenshot LSE006_MetadatenExport).

Metadatenkategorien: Keine Hinweise auf Webseite.

standardisierte Formate: Laut re3data Dublin Core DCC [re317b], aber keine Hinweise auf Webseite. "Metadata enhancements to record funding information eg, RIOXX, V4OA [...] Bibliographic metadata used to report to funders eg, HEFCE and EPrints REF 2014 Bazaar package for UK research reporting" [AM14].

unterstützende Tools: Einspeisen von Blog Posts: "A secondary part of the project has been working with the School's IT team to [...] auto-populate as many of the metadata fields as possible, which we are still working on. [...] the School, including a team in the Library, is currently implementing a Current Research Information System (CRIS), which we anticipate will significantly change existing repository workflows and reduce the amount of manual metadata entry required [...]" [Wil17, S. 4].

formularbasierte Eingabe: Metadaten werden von Bibliothekarx erstellt: "The metadata for these takes longer to create, both by virtue of dealing with the full text and due to the additional metadata requirements of recent years, such as funder and project details to comply with the RCUK policy on Open Access, date of acceptance, preferred Creative Commons attribution, etc., which meant that overall we were seeing a significant increase in the hours required to keep on top of this work." [Wil17, S. 3].

Hinweis auf Pflichtangaben: Keine Hinweise auf Webseite.

Hilfe bei Schlagworten (Systematik): Systematik für Feld Subject hinterlegt (siehe Screenshot LSE007_AdvancedSearch).

1.3 Dokumentation

Dokumentation der FD: Keine Hinweise auf Webseite.

Verlinkung auf Datenpublikation: Keine Hinweise auf Webseite.

K1.4 Zitierfähigkeit

Richtlinien zur Zitation von FD: Zitationsvorschlag auf Datensatzebene (siehe Screenshot

²⁸¹<http://www.lse.ac.uk/library/research-support/research-data-management/file-formats>

LSE001_AnzeigeDatensatz), sonst keine Hinweise auf Webseite.

PID: "LSE Library can create a DOI if an object meets the criteria of: - Not being owned by or hosted by a third party, like a publisher, and has an association with the school - for example a report published by an LSE department or research centre. - Are not given a DOI by someone else, like an academic publisher or data service. - Have a degree of stability. They are not likely to change rapidly or drastically. - The object has or will have a record in LSE Research Online or LSE Theses Online." [LSE Ja].

K1.5 Lizenzmodelle

Nutzungsbedingungen festlegen: "Full text publications placed in the repository are covered by copyright law and clear conditions of use are displayed on the web pages associated with the documents. [...] Readers will have no more rights to copy and download than in any other publication. Where available LSE Research Online contains links to an article's DOI and the homepage of the journal or publisher website, ensuring the published version is clearly identified to users of the repository." [LSE Je].

offene Lizenzen anbieten: Williams erwähnt das "preferred Creative Commons attribution" [Wil17, S. 4] ein Metadatum ist, keine Hinweise auf Webseite.

Datenschutz: Privacy and data protection Abschnitt in den Terms of Use der Universität [Lon J].

MD Bindungslos (CC0): Keine Hinweise auf Webseite.

K1.6 Verweise auf Veröffentlichungen

Link zu Veröffentlichung der Publikation: "Authors are encouraged to deposit work in pre-publication form, clearly labelled as such, alongside web references to publications. This will ensure your work is accessible to all but still closely associated with the publication that contains the published version." [LSE Je].

Link zu Publikationen die nachnutzen: Keine Hinweise auf Webseite.

Verknüpfung zugehöriger Projekte: Über Sammlungen / Collections (siehe Screenshot LSE001_AnzeigeDatensatz).

K1.7 Nutzungsstatistik

Metriken auf Datensatzebene: Zeigt Downloadzahlen seit Upload an (siehe Screenshot LSE002_Statistik). "We already collect statistics for publications downloaded from LSE Research Online, which can provide useful data alongside traditional citation counts." [Ayr13].

Standardisierte Erhebung: Keine Hinweise auf Webseite.

K1.8 Anerkennung der wissenschaftlichen Leistung

Anreize für Forscherx: "LSERO provides researchers with a platform to [...] establish their work among a global audience, by making it open access. Depositing research in LSERO increases its visibility and the potential for increased citations; it also creates opportunities for collaboration with other researchers. The RSS feeds feature allows visitors to keep up to date with research from a particular LSE author or department. [...] Why is it important for LSE researchers to showcase their work in this way? It is important because open access has been widely linked to generating positive outcomes in scholarship and career progress in modern academia [...]" [LSE14]. "[...] provide stable links to published items [...] research outputs included in LSE Research Online will appear in LSE authors' Experts profiles automatically [...] adding citations and full text versions aids in the dissemination of research." [LSE Je].

Intregation PID: Publikationen über BibTex in ORCID einspeisen: "You can easily import your citations from LSE Research Online to your ORCID record in just a few clicks. [...] Citations for everything from LSE Research Online will now appear in your ORCID record." [LSE Jc]. "Adopting a reliable author name authority eg, ORCID" [AM14]. Keine weiteren Hinweise auf Webseite.

Zitationsraten / Altmetrics: Altmetrics, siehe Screenshot LSE008_Altmetric. "More recently we have added an altmetrics tool to measure the ways in which LSE research is being used in this broader social landscape. The Altmetric badge, which is displayed on every item in LSE Research Online with a digital object identifier (DOI), displays a score of attention and other accompanying metrics to help researchers track online attention surrounding their work." [Ayr13].

automatische Reportings: Keine Hinweise auf Webseite.

K2.1 Zugangsformen

öffentlich / Nutzergruppen: Open und Restricted laut re3data [re317b] , keine weiteren Hinweise auf Webseite.

Nutzerkreis Abgabe: "We can accept research from LSE Academic and non-Academic staff members, as well as LSE PhD research students." [LSE Je].

Zugriffsteuerung (rechtliche / ethische Einschränkungen): Keine Hinweise auf Webseite.

Metadaten restriktionsfrei: Keine Hinweise auf Webseite.

K2.2 Embargofunktion

Embargofunktion: "There are currently 1,300 such documents in LSERO, with many embargoed for up to two years from publication." [Fla17].

Embargofristen: Keine Hinweise auf Webseite.

Begründung & Ende sichtbar: Keine Hinweise auf Webseite.

K2.3 Recherchemöglichkeiten

Allgemein: Browsing eingeschränkt möglich (siehe Screenshot LSE003_Browse).

Freie Textsuche: Einfache Einschlitzsuche und Erweiterte Suche mit diversen Auswahlmöglichkeiten, z.B. Freitextsuche in Documents, Title, Abstract und Uncontrolled Keywords, sowie Einschränkung auf Subjects, Item Type, Status usw. (siehe Screenshot LSE007_AdvancedSearch).

Suchoperatoren: Keine Suchhilfe. Test: Bool erfolgreich, Trunkierung, Wildcard ohne Fehlermeldung (siehe Screenshot LSE009_Trunkierung).

attributspezifisch (Metadaten): Nur einige, siehe LSE007_AdvancedSearch.

Facettierung: Keine Facettierung.

Empfehlungssysteme: Keine Hinweise auf Webseite.

MD-Suche / Inhaltliche Suche: Keine expliziten Aussagen auf Webseite, aber Suchfeld "Documents" impliziert Suche in Textdokumenten.

Indexierbare Landingpage: "[...] openly share its information with internal and external services, such as LSE Experts and Google." [LSE Je]. "The contents of LSERO are highly discoverable through Google and Google Scholar, meaning more people are able to find and access the important research carried out at LSE." [LF16].

2.4 Integration in Zitationsdatenbanken

Metadaten bereit / liefert an Zit.-DB: Keine Hinweise auf Webseite.

MD-Abfrage-Schnittstellen (OAI-PMH): Hat eine OAI-PMH Schnittstelle [LSE Jg].

K2.5 Programmierschnittstellen

APIs: "LSE Research Online supports the Open Archives Initiative Protocol for Metadata Harvesting." [LSE Je].

K3.1 Policies

Richtlinien: Mission Statement enthalten in FAQ²⁸². Takedown Policy²⁸³. LSE Open Access

²⁸² <http://eprints.lse.ac.uk/faq.html>

²⁸³ <http://eprints.lse.ac.uk/takedown.html>

Publications Policy²⁸⁴. Research Data Management Policy²⁸⁵. Allgemeine Nutzungsbedingungen der Webseite der Universität²⁸⁶.

K3.2 Datenvalidierung

"Bibliographic details will be checked by Library staff before a paper is made publicly available." [LSE Je].

K3.3 Versionierung

Versionierung: Kein Hinweis auf Webseite gefunden.

aktuelle V. gekennzeichnet / ältere V. verknüpft: Nicht zutreffend.

Änderungen protokolliert: Nicht zutreffend.

PID-Vergabe deutlich: Nicht zutreffend.

K3.4 Löschung

Geregelt: Nur in den OAI PMH Daten: "Deleted Record Policy: persistent" [LSE Jg]. Explizite Informationen zur Löschung sonst nur bei Verstößen gegen Copyright: "If it is concluded that the material is in breach of intellectual property rights, the material will be permanently withdrawn from the LSE Research Online website." [LSE Jf].

URL & PID funktionsfähig: "A citation/citations to the Work(s) will always remain visible in the repository during its lifetime." [LSE Je].

Grund der Löschung auf Landing-Page: Keine Hinweise auf Webseite.

K3.5 Sicherstellung langfristiger Zugänglichkeit

Übernimmt Verantwortung: "May electronically store, translate, copy, or re-arrange the Work(s) to ensure its/their future preservation and accessibility within the lifetime of the repository, that is, to ensure it can be read by computer system in the future [...]" [LSE Je].

Sicherungsmaßnahmen: "May electronically store, translate, copy, or re-arrange the Work(s) to ensure its/their future preservation and accessibility within the lifetime of the repository, that is, to ensure it can be read by computer system in the future [...]" [LSE Je].

Sicherungsdauer: Keine Hinweise auf Webseite.

Nachfolgeregelungen: Keine Hinweise auf Webseite.

K3.6 Zertifizierung

Nicht zertifiziert.

Einhaltung von Anforderungen: Nicht zutreffend.

Prüfberichte Online: Nicht zutreffend.

K4.1 Usability

Nutzerxfreundlichkeit: "You can easily import your citations from LSE Research Online to your ORCID record in just a few clicks." [LSE Jc]. "Adding repository content as data source to a unified resource discovery tool eg, Primo [...] Populating staff profile or webpages publication lists." [AM14].

intuitive Bedienung: Nein, z.B. keine leere Suche möglich um alle Suchergebnisse anzeigen zu lassen (siehe Screenshot LSE010_LeereSuche), keine Anpassung der Suchbegriffe auf Ergebnisseite (siehe Screenshot LSE004_AnzahlErgebnisse).

etablierte Darstellungsformen: Ja, aber veraltet.

flache Lernkurve: Nein, Tooltips nicht immer Aussagekräftig (siehe Screenshot LSE011_Tooltips), keine Suchhilfe.

Barrierefreiheit: Universität hat einen Accessibility Absatz in den Terms of Use: "LSE does

²⁸⁴ https://info.lse.ac.uk/staff/services/Policies-and-procedures/Assets/Documents/opeAccPub.pdf?from_serp=1

²⁸⁵ <https://info.lse.ac.uk/staff/services/Policies-and-procedures/Assets/Documents/resDatManPol.pdf>

²⁸⁶ <http://www.lse.ac.uk/lse-information/terms-of-use>

its best to ensure that the Website is accessible to all its users and potential users, including but not limited to users with disabilities (eg visual impairments, cognitive impairments, motor impairments etc); users for whom English is a second language; users of various software; or any other users who may need to use the Website. [...] If you need any advice on making your web content and services accessible, please contact the Web Services Team." [Lon J]. Browsererweiterung aXe findet 14 Verstöße gegen Barrierefreiheit (siehe Screenshot LSE012_Accessibility-Suche), fehlender Farbkontrast, fehlende Beschriftung von Form tags. 4 Verstöße auf Datensatzebene (siehe Screenshot LSE013_Accessibility-Datensatz), wieder fehlender Farbkontrast. Manuelle Betrachtung: Keine Hinweise zur Beschreibung von Videos. Keine Funktion für das Einstellen von Transkript oder Untertitel gefunden. Keine Textalternativen (Stichprobenprüfung für jeweils 3 Datensätze aus Item Type Video, Image, Audio). Keine Anpassung Textgröße oder Farbschema. Nicht mit Tastatur allein bedienbar. Keine zeitlichen Einschränkungen. Keine Blitze. Keine Bemühungen hinsichtlich Lesbarkeit. Keine Eingabehilfe oder Fehlerkorrektur in der Suche.

K4.2 Darstellung der Datensätze

Datenvisualisierung: Keine Hinweise auf Webseite.

Vorschaufunktion: Videos und Bilder im Browser.

K4.3 Support zu Repositoriumsfunktionen

Hilfeseiten: Nur FAQs²⁸⁷.

Schulungsangebote: Keine Hinweise auf Webseite.

Supportanfragen: Verweis auf eMail Service²⁸⁸.

Fehler melden: Keine Hinweise auf Webseite.

K4.4 Beratung zum Forschungsdatenmanagement

Beratung: Über die Bibliothek²⁸⁹. "The service is here to help you: - Understand expectations of LSE, research funders, and academic journals in sharing research data. - Manage research data from planning through to finishing a research project. - Describe data for use or re-use in your research. - Preserve and share data in a suitable data archive or repository with an appropriate re-use license." [LSE Jd], z.B. Data management planning, Dateiformate, Metadaten, Versionierung, Moodle Kurs zu RDM (nicht öffentlich)

Hilfestellung: "Send us an email to get advice and support about all aspects of data management." [LSE Jd]. Verweis auf aktuelle Richtlinien / Standards: "Find out your publisher's position using the RoMEO website or by checking the agreement you signed. The Library will help if you are not sure." [LSE Je]. Verweis auf UK Data Service Recommended Formats und Library of Congress Recommended Format Specifications [LSE Jb].

K5.1 Organisationsform

Passt zu Zielen / Aufgabe: Ansiedlung an der Bibliothek passend zu Ziel: "LSE Research Online aims to be a complete database of research produced by LSE Staff." [LSE Je].

Langfristigkeit: Keine Hinweise auf Webseite.

Träger bekannt: "LSE Library manages LSE Research Online on behalf of the School as a whole." [LSE Je].

Gemeinnützig: Keine Hinweise auf Webseite.

K5.2 Personalausstattung und -qualifizierung

"Library staff managing the repository are information professionals with experience of working with researchers and of creating high quality catalogue and index records." [LSE Je]. "The Me-

²⁸⁷ <http://eprints.lse.ac.uk/faq.html>

²⁸⁸ <http://eprints.lse.ac.uk/contact.html>

²⁸⁹ <http://www.lse.ac.uk/library/research-support/research-data-management>

tadata Services team currently consists of the Manager, 2 Assistant Librarians (one of whom is part-time), 2 Senior Library Assistants, and 3 Library Assistants. From this group of 8, half are involved in working on our LSE Research Online (LSERO) repository. One of our Senior Library Assistants oversees and manages repository work in Metadata Services, and our 3 Library Assistants combine work on LSE Research Online with cataloguing work." [Wil17].

K5.3 Forschungs- und Entwicklungstätigkeiten

Steuert Weiterentwicklung der Infrastruktur: Keine Hinweise auf Webseite.

Kooperationen mit FDM- Initiativen und Fachdisziplin: Keine Hinweise auf Webseite.

Mitarbeit an FDM-Standards: Keine Hinweise auf Webseite.

Forschungsprojekte: Keine Hinweise auf Webseite.

K5.4 Technische Infrastruktur

Allgemein: "This site is powered by EPrints 3." [LSE Je]. "Statistics packages e.g. EPrints IRStats2" [AM14].

Angemessen, Zuverlässig, Sicher, Skaliert: Keine Hinweise auf Webseite.

schnelle Lade- & Reaktionszeiten: Suche sehr langsam.

Ausfallsicherheit (verteilte DB, Server): Keine Hinweise auf Webseite.

Back-Up Strategie: Keine Hinweise auf Webseite.

Überwachung der Komponenten: Keine Hinweise auf Webseite.

techn. Infrastrukt. dokumentiert: Keine Hinweise auf Webseite.

offene Software: "This system is running eprints server software (EPrints 3.3.15) developed at the University of Southampton." [LSE Jg].

Forschungseth. und rechtl. Richtlinien: Keine Hinweise auf Webseite.

K6.1 Gebühren für die Datenabgabe und -nutzung

Bekannte Gebührenmodelle (Dauer / Volumen Staffellung): Keine Hinweise auf Webseite.

Verträge zwischen FDR und Institut: Keine Hinweise auf Webseite.

K6.2 Finanzierung des Betriebs

Langfristiges Konzept: Keine Hinweise auf Webseite.

Budgetplanung: Keine Hinweise auf Webseite.

Prognosen für Zukunft: Keine Hinweise auf Webseite.

Oceanrep GEOMAR Repository

Beispieldatensätze für Screenshots:

Beispiel 1: http://dx.doi.org/10.3289/MATING_CLIONE_LIMACINA_2010

Beispiel 2: http://dx.doi.org/10.3289/KOSMOS_SEDIMENT_TRAP_SAMPLING

re3data Datensatz für OceanRep [re316b].

K1.1 Datenabgabe

Allgemein: "[...] es gibt keine Begrenzung, was die Anzahl hochgeladener Dokumente betrifft." [GEO Ja]. "Die Einträge müssen folgende Kriterien erfüllen: Sie müssen für die Forschung in den marinen Naturwissenschaften relevant und so weit fertig gestellt sein, dass sie als Basis eines Kommunikationsprozesses dienen können. Außerdem müssen die Zusatzinformationen (Metadaten) korrekt angegeben werden. Es können Dokumente in allen Sprachen eingestellt werden [...]" [GEO Jd]. "In das Archiv können folgende Arten von Dokumenten eingestellt werden: Artikel aus Fachzeitschriften (peer-reviewed & non peer-reviewed), Artikel aus der allgemeinen Presse, Bücher, Buchkapitel, Konferenzbände, Artikel aus Konferenzbänden, Vorträge, Poster, Hochschulschriften (über Verlag veröffentlicht & nicht über Verlag veröffentlicht),

Lehrmaterialien, technische Berichte, interne Berichte, Patente, Videos, Tonaufnahmen, Bilder, Interviews/Fernsehauftitte und sonstige Dokumente.”[GEO Jd]

Zulässiges Format: "Die Dokumente sollen in einem einheitlichen, plattformunabhängigen Dokumentformat angeboten werden. Dieses Format ist derzeit PDF (Adobe Portable Document Format) [...] Bitte achten Sie darauf, dass Sie das PDF nicht sperren, da sonst eine Volltextindizierung nicht möglich ist. [...] Die folgenden Formate werden durchgängig von unserem System akzeptiert: PDF, Postscript, HTML, und ASCII." [GEO Ja]. Übersicht der Publikationstypen²⁹⁰.

Möglichkeiten des Uploads: "Wenn Sie eine Arbeit ins Repository einbringen möchten, haben Sie zwei Möglichkeiten: - Sie melden uns die Daten und wir geben sie ein. - Sie geben die Daten selber ein." [GEO Ja]. "Die Schaltfläche "Einträge importieren" erreichen Sie über die Rubrik "Einträge verwalten" in Ihrem Benutzer-Arbeitsbereich. Sie ermöglicht es Ihnen, Einträge aus anderen Quellen zu importieren und in einen Eprints-Eintrag zu überführen. Es können verschiedenen Formate verwendet werden: Pubmed ID, Pubmed XML, DOI system oder XML." [GEO Jb].

Daten-Integritätsprüfung: "Jeder Eintrag wird zunächst überprüft (Publikationsart, Pflichtfelder, Schreibfehler, Format und Urheberrecht) bevor er freigeschaltet wird." [GEO Ja].

K1.2 Metadaten

Metadatenkategorien: Keine Hinweise auf Webseite.

standardisierte Formate: Keine Hinweise auf Webseite.

unterstützende Tools: "Autoren: Geben Sie den Nachnamen entsprechend der Angaben auf Ihrem Dokument ein. Das System bietet automatisch Nach- und Vornamen an, nachdem die ersten Buchstaben eingegeben wurden. Sie können dann aus der Liste den entsprechenden Autoren auswählen. Es wird automatisch auch der Forschungsbereich ausgewählt." [GEO Jb].
formularbasierte Eingabe: "Der Prozess des Einstellens ist in fünf Schritte unterteilt. Zu jedem Schritt gibt es eine Eingabemaske (Typ, Details, Schlagworte, Datei hochladen, absenden)." [GEO Jb].

Hinweis auf Pflichtangaben: "Rote Sterne markieren Pflichtfelder." [GEO Jb].

Hilfe bei Schlagworten (Systematik): "Themengebiete wählen Sie durch Anklicken der Schaltfläche "Hinzufügen" rechts des jeweiligen Themengebietes aus. Wenn Sie auf das "+" links der Themengebiete klicken, werden weitere Unterkategorien angezeigt. [...] Wählen Sie mindestens ein Themengebiet und wahlweise weitere passende Themengebiete aus der Liste aus. Es wird empfohlen maximal drei Themengebiete auszuwählen. [...] Falls Sie der Meinung sind, dass wesentliche Themengebiete fehlen, wenden Sie sich bitte an die Bibliotheksleitung." [GEO Jb].

1.3 Dokumentation

Dokumentation der FD: Keine Hinweise auf Webseite.

Verlinkung auf Datenpublikation: "Offizielle URL: Falls Ihr Dokument an anderer Stelle online veröffentlicht wurde, können Sie hier die Internetadresse (URL, inkl. http://) angeben.[...]

Verwandte URLs: Falls Ihr Eintrag Teil eines übergeordneten, Kontextes ist, der online zur Verfügung steht, können Sie hier die entsprechenden Links (z.B. auf eine Projekt- oder Konferenzhomepage) angeben." [GEO Jb].

K1.4 Zitierfähigkeit

Richtlinien zur Zitation von FD: Zitationsvorschlag auf Datensatzebene (siehe Screenshot ORE001_AnzeigeDatensatz).

PID: Nutzt DOI (siehe Screenshot ORE001_AnzeigeDatensatz). "Außerdem archivieren wir Ihre wissenschaftlichen Publikationen dauerhaft und zitierfähig mit standardisierten Adressen (so genannte URNs), sodass Ihre Publikationen über (gewöhnliche) Suchmaschinen gefunden

²⁹⁰<http://oceanrep.geomar.de/view/types/public/>

werden." [GEO Ja].

K1.5 Lizenzmodelle

Nutzungsbedingungen festlegen: Lizenzen werden auf Datensatzeben angezeigt (siehe Screenshot ORE001_AnzeigeDatensatz). "Lizenz: Spezifiziert die Lizenz des Dokuments (Angaben, die Sie hier machen, haben keine Auswirkungen auf die Zugriffsrechte, die sie weiter oben definiert haben)." [GEO Jb].

offene Lizenzen anbieten: "Kann ich bestimmen, was in welchem rechtlichen Rahmen mit meiner Arbeit gemacht werden darf/kann? Ja, durch Lizenzangebote der gemeinnützigen Gesellschaft 'Creative Commons'" [GEO Ja]. "Lizenz: Spezifiziert die Lizenz des Dokuments (Angaben, die Sie hier machen, haben keine Auswirkungen auf die Zugriffsrechte, die sie weiter oben definiert haben). Dieses Online-Archiv ermöglicht die Auswahl zwischen verschiedenen CC-Lizenzen (so.g Creative Commons Licences). Falls keine der angegebenen Lizenzen für Ihr Dokument zutrifft, lassen Sie bei diesem Feld "Nicht spezifiziert" stehen." [GEO Jb].

Datenschutz: Keine Hinweise auf Webseite.

MD Bindungslos (CC0): "Anyone may access the metadata free of charge. The metadata may be re-used in any medium without prior permission for not-for-profit purposes provided the OAI Identifier or a link to the original metadata record are given." [GEO Jc].

K1.6 Verweise auf Veröffentlichungen

Link zu Veröffentlichung der Publikation: "Offizielle URL: Falls Ihr Dokument an anderer Stelle online veröffentlicht wurde, können Sie hier die Internetadresse (URL, inkl. http://) angeben." [GEO Jb]. Siehe Screenshot ORE004_Verlinkung.

Link zu Publikationen die nachnutzen: Keine Hinweise auf Webseite.

Verknüpfung zugehöriger Projekte: "Additional Affiliation: Hier ordnen Sie Ihren Eintrag dem passenden Forschungsbereich, der passenden Forschungseinheit bzw. dem Projekt zu, in dem das Dokument erstellt wurde. [...] Verwandte URLs: Falls Ihr Eintrag Teil eines übergeordneten, Kontextes ist, der online zur Verfügung steht, können Sie hier die entsprechenden Links (z.B. auf eine Projekt- oder Konferenzhomepage) angeben." [GEO Jb]. "Mittels des Institutional Repositories OceanRep wurde eine automatische Verknüpfung zu Projekten und Expeditionen (beziehungsweise Modellen und Experimenten), bei denen die Daten erhoben wurden, möglich. Dazu gehört eine automatische Verlinkung zum Datenmanagement-Portal [...] und eine automatische Verlinkung zu Pangaea [...], wenn dort Daten zu der Expedition liegen." [GEO15]. Verlinkung auf Projekte siehe Screenshot ORE013_VerlinkungProjekte.

K1.7 Nutzungsstatistik

Metriken auf Datensatzebene: Anzeige von Downloads pro Tag und Monat auf Datensatzebene siehe Screenshot ORE002_Statistik und erweiterte Anzeige inkl. ReferrerGraph, SearchEngineGraph, TopTenSearchTermsTable und MonthlyUniqueVisitorsGraph siehe Screenshot ORE003_MoreStatistics.

Standardisierte Erhebung: Keine Hinweise auf Webseite.

K1.8 Anerkennung der wissenschaftlichen Leistung

Anreize für Forscherx: "Untersuchungen zeigen, dass Open AccessPublikationen weitaus häufiger zitiert werden als Non Open Access Publikationen. [...] OceanRep wird darüber hinaus von verschiedenen Portalanbietern regelmäßig abgefragt. Damit erreichen Sie eine weitere Verbreitung Ihrer Arbeit und erhöhen Ihren persönlichen Impact-Faktor." [GEO Ja]. Anzeige der meist heruntergeladenen Paper²⁹¹. "Automatisches Update von Publikationslisten auf Projekt-, Schiffs-, und persönlichen Webseiten" [Sch14].

Intregation PID: "In OceanRep werden derzeit die GEOMAR-Mitarbeiter mit der GND, ResearcherID, ORCID und Scopus ID verknüpft - soweit vorhanden. In Zukunft wollen wir dann

²⁹¹http://oceanrep.geomar.de/cgi/irstats.cgi?page=dashboard&IRS_epchoice=All

in OceanRep die Möglichkeit bieten, aus einer ORCID ID Literatur in OceanRep hochzuladen, bzw. über die ORCID Einträge zu suchen." [Sch16].

Zitationsraten / Altmetrics: Keine Hinweise auf Webseite.

automatische Reportings: Anzeige der meist heruntergeladenen Paper²⁹². "Einfache und schnelle Auswertung für Jahresberichte und Evaluierung; es wurden neue Möglichkeiten der Listenerzeugung bereitgestellt, zum Beispiel für die LOM (Leistungsorientierte Mittelvergabe) sowie POF (Programmierorientierte Förderung)" ([Sch14, S. 59]).

K2.1 Zugangsformen

öffentlich / Nutzergruppen: "Es kann zwischen folgenden Kategorien gewählt werden: Jeder [...] Nur registrierte Benutzer [...] Nur Mitarbeiter des Archivs: Das Dokument ist nur für Sie selbst und für die Administratoren des Archivs zugänglich. Sie können diese Option z.B. für Quelldateien nutzen, die Sie nur zu Ihrem eigenen Gebrauch archivieren wollen. Weiterhin können so Veröffentlichungen archiviert werden, die strengen Copyrightbeschränkungen unterliegen. [...] Diese Option kann auch genutzt werden, um in Abstimmung mit einer Organisation Gutachtern einen begrenzten Zugriff auf nicht frei zugängliche Veröffentlichungen dieser Organisation zu ermöglichen." [GEO Jb].

Nutzerkreis Abgabe: "OceanRep ist ein Volltextarchiv für den wissenschaftlichen Output des GEOMAR und seiner MitarbeiterInnen." [GEO Ja]. "Items may only be deposited by accredited members of the organisation, or their delegated agents." [GEO Jc].

Zugriffsteuerung: "Nur Mitarbeiter des Archivs: Das Dokument ist nur für Sie selbst und für die Administratoren des Archivs zugänglich. Sie können diese Option z.B. für Quelldateien nutzen, die Sie nur zu Ihrem eigenen Gebrauch archivieren wollen. Weiterhin können so Veröffentlichungen archiviert werden, die strengen Copyrightbeschränkungen unterliegen. [...] Diese Option kann auch genutzt werden, um in Abstimmung mit einer Organisation Gutachtern einen begrenzten Zugriff auf nicht frei zugängliche Veröffentlichungen dieser Organisation zu ermöglichen." [GEO Jb].

Metadaten restriktionsfrei: Kontext Sichtbarkeit nur auf Mitarbeiter des Archivs beschränken: "Nutzer können die Metadaten einsehen (z.B. Titel, Autoren, Zusammenfassung), aber nicht das angehängte Dokument öffnen."

K2.2 Embargofunktion

Embargofunktion: Erweiterte Suche zeigt an: Zugangsrechte für das Dokument: Closed access, Embargoed access, Restricted access, Open access (siehe Screenshot ORE008_Zugangsrechte).

Embargofristen: "Enddatum der Sperre: Datum, an dem die durch einen Verlag oder einen Geldgeber verhängte Zugriffsbeschränkung ausläuft. Mit Erreichen des Datums wird das Dokument öffentlich zugänglich gemacht." [GEO Jb].

Begründung & Ende sichtbar: Keine Hinweise auf Webseite.

K2.3 Recherchemöglichkeiten

Allgemein: Freie Textsuche: "Die Schnellsuche ermöglicht eine rasche Durchsuchung der Felder, die i.d.R. am häufigsten durchsucht werden: 'Autor', 'Jahr', 'Titel', 'Zeitschrift' und 'Schlagworte'. [...] Das Archiv bietet neben der Schnellsuche zwei weitere Suchfunktionen an: die einfache und die erweiterte Suche." [GEO Jb].

Suchoperatoren: Bool AND und OR [GEO Jb]. Test: Trunkierung ohne Fehlermeldung, Wildcard nicht mit ? möglich.

attributspezifisch (Metadaten): "Das Blättern ermöglicht es Ihnen, das Archiv thematisch zu durchsuchen. [...] Wählen Sie in der Menüleiste am linken Rand des Bildschirms unter "Blättern" die Kategorie aus, nach der Sie das Archiv durchblättern wollen (Autor, Forschungsbereich, Publikationsart, Jahr, Themengebiet)." [GEO Jb].

²⁹² http://oceanrep.geomar.de/cgi/irstats.cgi?page=dashboard&IRS_epchoice=All

Facettierung: Keine Hinweise auf Webseite.

Empfehlungssysteme: Keine Hinweise auf Webseite.

MD-Suche / Inhaltliche Suche: "Bitte achten Sie darauf, dass Sie das PDF nicht sperren, da sonst eine Volltextindizierung nicht möglich ist." [GEO Jb]. Und in einfacher Suche "Alle Felder (einschließlich Volltexte)" weitere Hinweis auf inhaltliche Suche (siehe Screenshot ORE009_EinfacheSuche).

Indexierbare Landingpage: "Außerdem archivieren wir Ihre wissenschaftlichen Publikationen dauerhaft und zitierfähig mit standardisierten Adressen (so genannte URNs), sodass Ihre Publikationen über (gewöhnliche) Suchmaschinen gefunden werden." [GEO Ja].

2.4 Integration in Zitationsdatenbanken

Metadaten bereit / liefert an Zit.-DB: "Die Institutsbibliothek sorgt für den Austausch der Metadaten mit anderen Servern über OAI-PMH [...] Ja, die EPrints-Software kann die verwendeten Quellen und Angaben aus OceanRep in Programme, wie Endote exportieren." [GEO Ja].

MD-Abfrage-Schnittstellen (OAI-PMH): "OceanRep unterstützt OAI 2.0 mit der Basis-URL <http://oceanrep.geomar.de/cgi/oai2>." [GEOoJ].

K2.5 Programmierschnittstellen

APIs: Nur OAI-PMH (s.o.).

K3.1 Policies

Richtlinien: Eine explizite Richtlinie²⁹³. Institutsweite "Richtlinie für wissenschaftliche Veröffentlichungen" ist nur im Intranet verfügbar²⁹⁴.

K3.2 Datenvalidierung

"Die Dokumente werden von uns nach formaler und rechtlicher Prüfung zur Veröffentlichung freigegeben. [...] Jeder Eintrag wird zunächst überprüft (Publikationsart, Pflichtfelder, Schreibfehler, Format und Urheberrecht) bevor er freigeschaltet wird." [GEO Ja]. "The validity and authenticity of the content of submissions is the sole responsibility of the depositor." [GEO Jc].

"Der Eintrag wird von einem Bibliotheksmitarbeiter geprüft und anschließend in das öffentlich zugängliche Archiv eingestellt. Falls es irgendwelche Probleme mit dem Eintrag gibt (z.B. Inhaltlich nicht passend, fehlende Angaben, Probleme mit dem angehängten Dokument), wird der zuständige Bibliotheksmitarbeiter Kontakt zu Ihnen aufnehmen." [GEO Jb].

K3.3 Versionierung

Versionierung: "If necessary, an updated version may be deposited." [GEO Jc]. "Stattdessen ermöglicht das Archiv, aktuelle Versionen eines Eintrags einzustellen." [GEO Jd]. "Neue Version: Diese Option kann genutzt werden, um eine neue Version eines bestehenden Eintrags zu erstellen. Es wird eine exakte Kopie des bestehenden Eintrags erstellt, an der Sie anschließend Änderungen vornehmen können." [GEO Jb].

aktuelle V. gekennzeichnet / ältere V. verknüpft: "Nutzer werden automatisch zur jeweils aktuellsten Version geleitet." [GEO Jd]. Änderungen protokolliert: Keine Hinweise auf Webseite.

PID-Vergabe deutlich: "Die neue Version wird mit der alten Version verlinkt, befinden sich beide Versionen im öffentlich zugänglichen Archiv, werden Nutzer automatisch zu der neuen Version geleitet." [GEO Jb].

K3.4 Löschung

Geregelt: "Kann ich ein schon veröffentlichtes ePrint wieder entfernen? Nein, Sie können schon veröffentlichte Texte aber sperren lassen, indem Sie eine Anfrage an unser OceanRep-Team schicken." [GEO Ja]. "Acceptable reasons for withdrawal include: - Proven copyright violation or plagiarism, - Legal requirements and proven violations, - National Security, - Falsified

²⁹³ <http://oceanrep.geomar.de/policies.html>

²⁹⁴ <https://www.geomar.de/zentrum/einrichtungen/bibliothek/publikationen-at-geomar/>

research. Withdrawn items are not deleted per se, but are removed from public view." citep-RepositoryPolicies.o.J.. "Grundsätzlich werden bereits veröffentlichte Einträge nicht aus dem Archiv entfernt, da dadurch Verlinkungen auf den Eintrag ins Leere führen. [...] Auf besonderen Wunsch kann der Archivadministrator Einträge sperren." [GEO Jd]. "Falls Sie Einträge, die Sie abgesendet haben oder Einträge aus dem öffentlich zugänglichen Archiv löschen wollen, wenden Sie sich bitte unter Angabe des Grundes an die Bibliotheksleitung." [GEO Jb].

URL & PID funktionsfähig: "Withdrawn items' identifiers/URLs are retained indefinitely. URLs will continue to point to 'tombstone' citations, to avoid broken links and to retain item histories." [GEO Jc].

Grund der Löschung auf Landing-Page: Kein Hinweis zu Anzeige, aber Grund muss angegeben werden: "Falls Sie Einträge, die Sie abgesendet haben oder Einträge aus dem öffentlich zugänglichen Archiv löschen wollen, wenden Sie sich bitte unter Angabe des Grundes an die Bibliotheksleitung." [GEO Jb].

K3.5 Sicherstellung langfristiger Zugänglichkeit

Übernimmt Verantwortung: Abschnitt Preservation Policy in den Repository Policies²⁹⁵.

Sicherungsmaßnahmen: "The repository will try to ensure continued readability and accessibility. The repository regularly backs up its files according to current best practice." [GEO Jc].

Allgemein: "Sie räumen uns das nicht-ausschließliche Recht (=einfaches Nutzungsrecht) ein, Ihre Arbeit auf OceanRep zu veröffentlichen, zu archivieren und zur Sicherung der Langzeitverfügbarkeit ggf. technische Veränderungen vorzunehmen." [GEO Ja] Sicherungsdauer: "Die Dokumente werden langfristig unter einer verlässlichen Struktur auf dem OceanRep-Server gespeichert. Die Verfügbarkeit soll so lange wie möglich gewährleistet werden." [GEO Ja]. "Items will be retained indefinitely." [GEO Jc].

Nachfolgeregelungen: "In the event of the repository being closed down, the database will be transferred to another appropriate archive." [GEO Jc].

K3.6 Zertifizierung

Nicht zertifiziert.

Einhaltung von Anforderungen: Nicht zutreffend.

Prüfberichte Online: Nicht zutreffend.

K4.1 Usability

Nutzerxfreundlichkeit: Diverse Exportformate (siehe Screenshot ORE005_Exportformate). "Automatisches Update von Publikationslisten auf Projekt-, Schiffs-, und persönlichen Webseiten" [Sch14]. "Als Vorlage verwenden: Diese Option kann genutzt werden, um einen bestehenden Eintrag als Vorlage für einen neuen Eintrag zu verwenden. [...] Sie haben die Möglichkeit, den Prozess des Einstellens zu unterbrechen. Auf jeder der fünf Eingabemasken gibt es die Schaltfläche "Für später speichern". Wenn Sie diese anklicken, wird der Eintrag gespeichert und Sie können ihn zu einem späteren Zeitpunkt weiter bearbeiten. [...] Typ des Eintrags: Wählen Sie die Option aus der Liste, die das, was Sie archivieren wollen, am besten beschreibt. Die Auswahl hat Einfluss darauf, welche Informationen auf den folgenden Seiten abgefragt werden. [...] Falls Fehler aufgetreten sind oder Informationen fehlen, wird Ihnen dieses in einem orange gefärbten Kasten am Anfang der Seite angezeigt." [GEO Jb].

intuitive Bedienung: Keine leere Suche möglich um alle Datensätze anzuzeigen. Keine Anpassung der Suchbegriffe auf Ergebnisseite (siehe Screenshot ORE014_SucheAnpassen).

etablierte Darstellungsformen: Ja, aber eher veraltet. Selbes Problem mit LSERO (auch eprints).

flache Lernkurve: "Rechts von jedem Eingabefeld befindet sich ein "?". Wenn Sie dieses anklicken, erhalten Sie Hilfestellung, was bei dem jeweiligen Feld anzugeben ist." [GEO Jb].

²⁹⁵ <http://oceanrep.geomar.de/policies.html>

Barrierefreiheit: Browsererweiterung aXe findet 49 Verstöße gegen Barrierefreiheit (siehe Screenshot ORE010_Accessibility-Suche), fehlender Farbkontrast, fehlende Beschriftung von Bildern und Links. 16 Verstöße auf Datensatzebene (siehe Screenshot ORE011_Accessibility-Datensatz), wieder fehlender Farbkontrast und Beschriftung von Bildern. Manuelle Betrachtung: Beschreibung von Videos kann im Abstract erfolgen, siehe Beispieldatensatz²⁹⁶ (Abstract entspricht bei Metadatenexport in DC dem dc:description Feld), aber keine spezifischen Metadatenfelder für inhaltliche Beschreibung, Transkript oder Untertitel gefunden. Keine Textalternativen (Stichprobenprüfung für 3 Datensätze aus Publikationsart Videos, keine "Volltext" Tonaufnahmen, keine Einträge in Bild). Auf Webseite des Geomar Zentrums ist Schriftgröße einstellbar (siehe Screenshot ORE007_Accessibility-Schrift), auf den Unterseiten des OceanRep ist keine Anpassung von Textgröße oder Farbschema möglich. Mit Tastatur allein bedienbar, aber schwierig. Keine zeitlichen Einschränkungen. Keine Blitze. Keine Bemühungen hinsichtlich Lesbarkeit. Keine Eingabehilfe oder Fehlerkorrektur in der Suche.

K4.2 Darstellung der Datensätze

Datenvisualisierung: Keine Hinweise auf Webseite.

Vorschaufunktion: "Auf dem Bildschirm, auf dem die Suchergebnisse angezeigt werden, können Sie: [...] Direkt die Dateien öffnen, die an dem Eintrag hängen." [GEO Jb]. Vorschau auch für Videos, Wiedergabe im Browser (siehe Screenshot ORE012_Vorschau).

K4.3 Support zu Repositoriumsaktionen

Hilfeseiten: Suchhilfe und Hilfeseiten für Forscherx²⁹⁷. FAQs²⁹⁸.

Schulungsangebote: Keine Hinweise auf Webseite.

Supportanfragen: "Wenn Sie unsicher sind, was Sie bei einem Feld eingeben sollen, klicken Sie auf das Fragezeichen rechts von dem entsprechenden Eingabefeld oder wenden Sie sich an die Bibliotheksmitarbeiter." [GEO Jb].

Fehler melden: Keine Hinweise auf Webseite.

K4.4 Beratung zum Forschungsdatenmanagement

Beratung: "Bitte kontaktieren Sie die Bibliothek, wenn Sie weitere Informationen oder Hilfe bezüglich der geeigneten Version des Materials benötigen." [GEO Ja]. "Falls Sie auf der Hilfeseite keine Antworten auf Ihre Fragen finden, wenden Sie sich bitte an die Bibliotheksleitung." [GEO Jb].

Hilfestellung: "Der Service, dass die Wissenschaftler ihre Meldungen einfach per E-Mail mit wenig Aufwand an die Bibliothek senden können und dass die Bibliothek für sie die Aufnahmen und Ergänzungen der Angaben übernimmt, wird sehr hoch geschätzt." [Sch14] (S. 60).

Verweis auf aktuelle Richtlinien / Standards: "Eine Übersicht über Richtlinien zum Umgang mit Forschungsdaten finden Sie unter 'Forschungsdaten.org!'" [GEO15].

K5.1 Organisationsform

Passt zu Zielen / Aufgabe: "Die Bibliothek ist damit ihrem Ziel einen großen Schritt näher gekommen, dass die Wissenschaftler bei dem Thema 'Publizieren' sofort an die Bibliothek als Ansprechpartner denken." [Sch14] (S. 60)

Langfristigkeit: Mitfinanziert und genutzt vom Exzellenz-Cluster 'Future Ocean' und Sonderforschungsbereichen 574 "Volatile und Fluide in Subduktionszonen" und 754 "Klima-Biogeochemische Wechselwirkungen im tropischen Ozean".

Träger bekannt: "OceanRep wird redaktionell von der GEOMAR-Bibliothek und technisch vom Kieler Datenmanagementteam betreut und wurde mitfinanziert vom Exzellenz-Cluster 'Future

²⁹⁶ http://dx.doi.org/10.3289/MATING_CLIONE_LIMACINA_2010

²⁹⁷ <http://oceanrep.geomar.de/help>

²⁹⁸ <http://oceanrep.geomar.de/faqs.html>

Ocean!" [GEOoJ]. "Der Cluster sowie beide Sonderforschungsbereiche entschieden sich zur Beteiligung an dem Repository." ([Sch14, S. 60]), gemeint sind die Sonderforschungsbereiche 574 "Volatile und Fluide in Subduktionszonen" und 754 "Klima-Biogeochemische Wechselwirkungen im tropischen Ozean".

Gemeinnützig: "Das GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel ist als Stiftung öffentlichen Rechts der Förderung der wissenschaftlichen Forschung auf dem Gebiet der marinen Naturwissenschaften verpflichtet." [Sch14] (S. 58).

K5.2 Personalausstattung und -qualifizierung

Die Bibliothek hat 3 Mitarbeiter²⁹⁹. Das Kieler Datenmanagement-Team (KDMT) hat 5 Mitarbeiter³⁰⁰. Keine Hinweise wer an OceanRep mitarbeitet und in welcher Kapazität.

K5.3 Forschungs- und Entwicklungstätigkeiten

Steuert Weiterentwicklung der Infrastruktur: Keine Hinweise auf Webseite.

Kooperationen mit FDM- Initiativen und Fachdisziplin: Keine Hinweise auf Webseite.

Mitarbeit an FDM-Standards: Keine Hinweise auf Webseite.

Forschungsprojekte: Bibliothek ist beteiligt an den Projekten PubFlow und MaNIDA³⁰¹.

K5.4 Technische Infrastruktur

Angemessen, Zuverlässig, Sicher, Skaliert: "EPrints wurde von Wissenschaftlern aus Southampton erarbeitet und dort von einer inzwischen festen Arbeitsgruppe von Wissenschaftlern betreut und weiterentwickelt. Der große Vorteil von EPrints ist, dass die Software von Wissenschaftlern für Wissenschaftler erstellt wurde und dass sie nicht von einer kommerziellen Firma vertrieben wird." [Sch14] (S. 57).

schnelle Lade- & Reaktionszeiten: Eher langsam, vor allem Suche.

Ausfallsicherheit (verteilte DB, Server): Keine Hinweise auf Webseite.

Back-Up Strategie: "The repository regularly backs up its files according to current best practice." [GEO Jc].

Überwachung der Komponenten: Keine Hinweise auf Webseite.

techn. Infrastrukt. dokumentiert: Keine Hinweise auf Webseite.

offene Software: EPrints.

Forschungseth. und rechtl. Richtlinien: Keine Hinweise auf Webseite.

K6.1 Gebühren für die Datenabgabe und -nutzung

Bekannte Gebührenmodelle: "Anyone may access full items free of charge." [GEO Jc].

Verträge zwischen FDR und Institut: Keine Hinweise auf Webseite.

K6.2 Finanzierung des Betriebs

Langfristiges Konzept: Keine Hinweise auf Webseite.

Budgetplanung: "An laufenden Kosten für Wartung und Updates ist mit 3000 Euro pro Jahr zu rechnen." [Sch14] (S. 60).

Prognosen für Zukunft: Keine Hinweise auf Webseite.

Marine-Geo Digital Library

Beispieldatensätze für Screenshots:

Beispiel 1: http://www.marine-geo.org/tools/search/Files.php?data_set_uid=24210

²⁹⁹ <https://www.geomar.de/zentrum/einrichtungen/bibliothek/team/>

³⁰⁰ <https://www.geomar.de/de/zentrum/einrichtungen/rz/daten/>

³⁰¹ <https://www.geomar.de/zentrum/einrichtungen/bibliothek/a-z/projekte/#c12388>

bzw. <http://get.iedadata.org/doi/324210> (unterschiedliche Anzeigen)

Beispiel 2: http://www.marine-geo.org/tools/search/Files.php?data_set_uid=17126

bzw. <http://get.iedadata.org/doi/317126> (unterschiedliche Anzeigen)

Beispiel 3: http://www.marine-geo.org/tools/search/Files.php?data_set_uid=24121

(Nur Embargo)

re3data Datensatz für die MGDL [[re317c](#)] (verzeichnet als "Marine Geoscience Data System").

K1.1 Datenabgabe

Allgemein: "The Marine-Geo Digital Library handles a wide range of geoscience data types including geophysical data, bathymetry, side-scan sonar, imagery, maps, seismic reflection, gridded compilations, earthquake catalogs, and other derived data products." [[Mar Jb](#)].

Zulässiges Format: "MGDS can accomodate a variety of file types, many of which are self-describing." [[Mar Jb](#)]. Teil des MGDS Vocabulary ist eine Liste der Formate³⁰², für Video wird .mov empfohlen.

Möglichkeiten des Uploads: "An integrated IEDA Data Submission interface is currently under development. Until it is released, please contribute data using our Data Registration Form. We can copy your data files from the file-sharing platform of your choice (e.g. DropBox, Google Drive, FTP). [...] Or, Login to GeoPass to use the legacy File Uploader tool." [[Mar Jd](#)]. Nach Login, auf der selben URL: "Our File Uploader enables the upload of data files for inclusion in the Marine Geo Digital Library. You may upload multiple files per submission, but uploads are limited to a single file format and 100 MB per session, 50 MB per file." [[Mar Jd](#)] aber nur für Excel spreadsheets, , PDF, NetCDF grid files und Shapefiles, nicht Video.

Daten- Integritätsprüfung: Keine Hinweise auf Webseite.

K1.2 Metadaten

Metadatenkategorien: "For each data object, we catalog basic metadata (file name, format, size etc), spatial/temporal information (optional, but recommended), and URLs to remote data centers (e.g. GenBank, SESAR). The database includes custom child tables for accommodating metadata for special file types (e.g. image, seismic image, grid, geotiff, seismic, swath). [...] Additional metadata that is cataloged to describe Expeditions/Compilations includes: Geographic/Temporal Bounds, Technical Reports, Scientific References, Data Management Plans, Related Web Pages, Funding Information, Participants and their Affiliations." [[Mar Jf](#)].

standardisierte Formate: Abfrage über Schnittstellen gibt ISO19115 Metadaten aus [[Mar Jh](#)]. Deutsche Übersetzung von ISO19115 - Geographic Information auf Geoportal.de³⁰³.

unterstützende Tools: "Auto-complete functionality helps fill in many of the fields - start typing and select one item from the corresponding options that pop up. The Select From List link brings up all options." [[Mar Jc](#)].

formularbasierte Eingabe: Data Registration Form auf Google Docs³⁰⁴ und Marine-Geo Digital Library File Uploader (kostenfreie Registrierung nötig). File Uploader siehe Screenshot MGD003_UploadForm.

Hinweis auf Pflichtangaben: Data Registration Form auf Google Docs, enthält Hinweise zu Pflichtangaben, z.B. Data Types, Device Type, Make, Model, Data Processing Level, Data Release Date. "The basic metadata we request includes: Names of scientists to whom the data should be credited, References in which the data were published, Related field program information, Data type and device details, Funding agency and award numbers." [[Mar Jb](#)]. Marine-Geo

³⁰²http://www.marine-geo.org/tools/search/vocab.php?use_is_displayed=T&vocab=vocab_file_format_id

³⁰³http://www.geoportal.de/SharedDocs/Downloads/DE/GDI-DE/Deutsche_Uebersetzung_der_ISO-Felder.html

³⁰⁴<https://docs.google.com/forms/d/e/1FAIpQLSf5Uu0eBZ71rjb29WivA3vFuFHYy29DobxLWXFSM0fCbJX0iw/viewform>

Digital Library File Uploader zeigt Pflichtangaben ebenfalls an, z.B. Investigator(s), Data Type(s), Expedition or Compilation ID(s) (siehe Screenshot MGD003_UploadForm).

Hilfe bei Schlagworten (Systematik): Keine Hinweise auf Webseite.

1.3 Dokumentation

Dokumentation der FD: Keine Hinweise auf Webseite, verweisen aber auf den Data Management Plan der National Science Foundation in denen Dokumentation gefordert wird: "All proposals must include a Data Management Plan that describes how the project will provide open and rapid access to quality-controlled and fully documented data and information during and after the project. This plan must be included as a Supplementary Document and be consistent with NSF's policy on dissemination and sharing of research results." [Nat16].

Verlinkung auf Datenpublikation: Möglich über "Related Reference(s)" Metadatum im Marine-Geo Digital Library File Uploader (siehe Screenshot MGD003_UploadForm). "It not only provides access to data files of a variety of formats and data types, but it enables data discovery by providing a comprehensive set of metadata including information about field operations and data acquisition, as well as data provenance (linking derived data to original raw data files), and scientific publications." [Mar Jf]

K1.4 Zitierfähigkeit

Richtlinien zur Zitation von FD: Zitationsrichtlinie in den Terms of Use für die einzelnen Portale [Mar Jj]. Anzeige auf Datensatzebene im Reiter "Data Citation Information" (siehe Screenshot MGD006_Citation).

PID: DOI, siehe Screenshot MGD005_Datensatzansicht. "Data submitted to the Marine-Geo Digital Library can be registered with DOIs[R] for persistent identification and formal citation." [Mar Je].

K1.5 Lizenzmodelle

Nutzungsbedingungen festlegen: "Appropriate acknowledgment with a byline/credit/link must be given to both the original scientists/data contributors by reference to their relevant publications and to the Marine Geoscience Data System (www.marine-geo.org). Where citation information has been provided to us by scientists it is included with the relevant database entries, and should be acknowledged when data are used. You may browse freely, but you may not circulate or publish materials you obtained from this site if you do not accept the terms of providing adequate citation." [Mar Jj].

offene Lizenzen anbieten: "Data and metadata presented on these pages is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License." [Mar Ja].

Datenschutz: Kein Hinweis auf Webseite.

MD Bindungslos (CC0): Nein, CC-BY-NC-SA: "Data and metadata presented on these pages is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License." [Mar Ja].

K1.6 Verweise auf Veröffentlichungen

Link zu Veröffentlichung der Publikation: "It not only provides access to data files of a variety of formats and data types, but it enables data discovery by providing a comprehensive set of metadata including information about field operations and data acquisition, as well as data provenance (linking derived data to original raw data files), and scientific publications." [Mar Jf]. Mindestens für die zugehörige Publikation, siehe Screenshot MGD007_References. "Through collaborations with several earth-science journals, we provide links between publications and data." [Mar Je].

Link zu Publikationen die nachnutzen: Keine Hinweise auf Webseite.

Verknüpfung zugehöriger Projekte: "In addition, available information about funding source(s),

related projects, related data and information at external websites is also provided." [Mar Jf].
Für Datensätze siehe Screenshot MGD008_RelatedDataSets.

K1.7 Nutzungsstatistik

Metriken auf Datensatzebene: Keine Anzeige auf Webseite, aber scheinbar Erhebung für Autorx. Siehe K1.8.

Standardisierte Erhebung: Keine Hinweise auf Webseite.

Allgemein: MGDS zeichnet diverse Statistiken auf Seitenebene auf, Übersicht auf Webseite³⁰⁵, darunter auch gesamte Zitationen der FD³⁰⁶.

K1.8 Anerkennung der wissenschaftlichen Leistung

Anreize für Forscherx: Keine Hinweise auf Webseite.

Intregation PID: Keine Hinweise auf Webseite.

Zitationsraten / Altmetrics: Keine Anzeige auf Webseite, aber scheinbar Erhebung für Autorx.

automatische Reportings: "We promote the collegial use of data through its Terms of Use and provides [sic!] semi-annual data download reports to data contributors to inform them of downloads/views of their data." [Mar Je]. "Data Curation Services include: [...] - Bi-annual data download & usage reports." [Mar Jd].

K2.1 Zugangsformen

öffentlich / Nutzergruppen: "The site contains both public and proprietary data. Generally, access to proprietary data is by mutual agreement and depends on terms and conditions of the original acquisition project." [Mar Jj].

"Data can be maintained within the system with proprietary hold or password protection to ensure preservation and promote collaboration prior to public release." [Mar Jd] nach Login.

Nutzerkreis Abgabe: "We welcome contributions of a broad range of sensor-based geoscience data." [Mar Jd] vor Login. "The Marine-Geo Digital Library hosts primarily marine geoscience data from the seafloor and subseafloor acquired with ships, towed platforms and submersibles and includes both derived data products and supporting field data. [...] The MGDS encourages submissions of new marine geoscience data sets for inclusion in the Marine-Geo Digital Library, in particular processed and interpreted data products supporting publications and field data that has not been archived." [Mar Je].

Zugriffsteuerung (rechtliche / ethische Einschränkungen): "Data can be maintained within the system with proprietary hold or password protection to ensure preservation and promote collaboration prior to public release." [Mar Jd] nach Login.

Metadaten restriktionsfrei: n.a.

K2.2 Embargofunktion

Embargofunktion: "Data can be maintained within the system with proprietary hold or password protection to ensure preservation and promote collaboration prior to public release." [Mar Jd] nach Login.

Embargofristen: Keine Hinweise auf Webseite.

Begründung & Ende sichtbar: Keine Begründung, aber Frist auf Datensatzebene sichtbar (siehe Screenshot MGD009_Embargo).

K2.3 Recherchemöglichkeiten

Allgemein: Suche in MGDL direkt oder über IEDA Integrated Catalog (Beta Version)³⁰⁷.

"MGDS provides Community Data Portals for several NSF-supported research programs and facilities based on disciplinary and programmatic themes. Each portal provides custom resour-

³⁰⁵ http://www.marine-geo.org/about/statistics_holdings.php

³⁰⁶ <http://www.marine-geo.org/about/citations.php>

³⁰⁷ <http://catalog-dev.iedadata.org:8080/geoportals/>

ces and access to comprehensive data collections assembled for each community." [Mar Jg].
 Freie Textsuche: Über den IEDA Integrated Data Browser ist Freitextsuche möglich (siehe Screenshot MGD011_IEDAIntegratedDataBrowser), direkt im FDR nur über Auswahllisten.
 Suchoperatoren: Keine auf Portal, da keine Freitextsuche.
 attributspezifisch (Metadaten): "The search interfaces enable searching based on one or more parameters such as data type, scientist name, location, citation." [Mar Ji].
 Suche kann über eine Karte (siehe Screenshot MGD010_Kartensuche) oder Textanzeige begonnen werden. Beide: Diverse Felder, alle mit Vokabular verknüpft. Liste aller Vokabulare unter Vocabulary Services³⁰⁸, siehe Screenshot MGD012_Vokabulare.
 Facettierung: Über IEDA Integrated Data Browser wird Facettierung angeboten (siehe Screenshot MGD011_IEDAIntegratedDataBrowser), keine Hinweise im FDR direkt.
 Empfehlungssysteme: Keine Hinweise auf Webseite.
 MD-Suche / Inhaltliche Suche: Nur MD-Suche, keine anderen Hinweise oder Eingabemöglichkeiten.
 Indexierbare Landingpage: Keine Hinweise auf Webseite. Known-Item Search (Leventer Near-bottom Videos Southern Ocean Palmer expedition 2014) für Beispieldatensatz <http://doi.org/10.1594/IEDA/324210> nicht erfolgreich.

2.4 Integration in Zitationsdatenbanken

Metadaten bereit / liefert an Zit.-DB: Keine Hinweise auf Webseite.
 MD-Abfrage-Schnittstellen (OAI-PMH): "MGDS FileServer provides access to file and data set level metadata from the Marine-Geo Digital Library including detailed file boundary geometries." [Mar Jh].

K2.5 Programmierschnittstellen

APIs: "We offer simple REST and OGC services to directly query our Digital Library for data and metadata. [...] We offer an API for querying for file metadata and downloading files and tar bundles directly from the MGD. [...] Marine-Geo Digital Library Search API exposes our search interface backend." [Mar Jh]. MGDS FileServer für Metadaten, findet die "data files identifiers" die der FileDownloadServer als Input benötigt. MGDS FileDownloadServer kann anhand der ID einzelne Dateien herunterladen. MGDS ArchiveDownloadServer kann bis max. 10 volle Datensätze herunterladen (Quelle für alle [Mar Jh])

K3.1 Policies

Richtlinien: Terms of Use³⁰⁹ und Data Submission Guidelines³¹⁰.

K3.2 Datenvalidierung

Keine Hinweise auf Webseite.

K3.3 Versionierung

Versionierung: Keine Hinweise auf Webseite.
 aktuelle V. gekennzeichnet / ältere V. verknüpft: Nicht zutreffend.
 Änderungen protokolliert: Nicht zutreffend.
 PID-Vergabe deutlich: Nicht zutreffend.

K3.4 Löschung

Geregelt: Keine Hinweise auf Webseite.
 URL & PID funktionsfähig: Nicht zutreffend.
 Grund der Löschung auf Landing-Page: Nicht zutreffend.

K3.5 Sicherstellung langfristiger Zugänglichkeit

³⁰⁸ http://www.marine-geo.org/tools/web_services.php

³⁰⁹ http://www.marine-geo.org/about/terms_of_use.php

³¹⁰ <http://www.marine-geo.org/submit/guidelines.php>

Übernimmt Verantwortung: "The Marine-Geo Digital Library offers data preservation services including long-term archiving, the registration of data with DOIs[R] for persistent identification and formal citation, and links to related scholarly publications." [Mar Jg].

Sicherungsmaßnahmen: Keine Hinweise auf Webseite.

Sicherungsdauer: Keine Hinweise auf Webseite.

Nachfolgeregelungen: Keine Hinweise auf Webseite.

K3.6 Zertifizierung

Nicht zertifiziert.

Einhaltung von Anforderungen: Nicht zutreffend.

Prüfberichte Online: Nicht zutreffend.

K4.1 Usability

Nutzerxfreundlichkeit: Suche auf Karte oder "text-only interface (for slow connections)" (siehe Screenshot MGD010_Kartensuche). News zeigt neueste Einträge in der MGDL³¹¹.

intuitive Bedienung: Textsuche eher ja, Kartensuche eher nein.

etablierte Darstellungsformen: Ja.

flache Lernkurve: Eher nein, nur eingeschränkte Suchhilfe.

Barrierefreiheit: Browsererweiterung aXe findet 32 Verstöße gegen Barrierefreiheit (siehe Screenshot MGD013_Accessibility-Suche), fehlender Farbkontrast, fehlende Beschriftung von Bildern, id-attribute mehrfach vergeben. 15 Verstöße auf Datensatzebene (siehe Screenshot MGD014_Accessibility-Datensatz), wieder fehlender Farbkontrast und Beschriftung von Bildern. Manuelle Betrachtung: Keine inhaltliche Beschreibung von Videos vorgesehen. Keine Funktionen für Transkript oder Untertitel gefunden. Keine Textalternativen (Stichprobenprüfung für 3 Datensätze aus Data Type Photograph, beinhaltet Video, keine Tondokumente). Nicht mit Tastatur allein bedienbar. Keine zeitlichen Einschränkungen. Keine Blitze. Keine Bemühungen hinsichtlich Lesbarkeit. Eingabehilfe vorhanden: "Each search parameter features case-independent auto-complete functionality" [Mar Jh]. Keine Fehlerkorrektur in der Suche.

K4.2 Darstellung der Datensätze

Datenvisualisierung: "A suite of portals provides custom tools for accessing and visualizing global collections of specialized data sets including gravity and magnetics trackline data, seismic reflection profiles, earthquake locations, seafloor bedrock geochemistry from PetDB, and sediment information from ocean drilling. Options are also available for users to import their own data tables and grids and make custom maps. Other visualization tools include Virtual Ocean, a 3D virtual globe application which integrates the GeoMapApp tool suite with the NASA World Wind 3-D earth browser and the IEDA Data Browser." [Mar Jg].

Vorschaufunktion: Siehe Screenshot MGD015_Vorschau für Bilder (Beispieldatensatz: http://www.marine-geo.org/tools/search/Files.php?data_set_uid=9890), keine Vorschaufunktion für Videos.

K4.3 Support zu Repositoriumsaktionen

Hilfeseiten: Zusammenstellung Frequently-Asked Questions about data terminology³¹², Suchhilfen³¹³, Übersicht über Programmierschnittstellen³¹⁴ und File Uploader - Tips and Help³¹⁵. Schulungsangebote: Videotutorials und Webinar für unterschiedliche Funktionalitäten³¹⁶.

³¹¹<http://www.marine-geo.org/portals/antarctic/news.php>

³¹²http://www.marine-geo.org/help/data_FAQ.php

³¹³http://www.marine-geo.org/help/search_results_help.php und http://www.marine-geo.org/help/search_results_help.php

³¹⁴http://www.marine-geo.org/tools/web_services.php

³¹⁵http://www.marine-geo.org/help/contribute_help.php

³¹⁶<http://www.marine-geo.org/help/index.php>

Supportanfragen: Über ein Kontaktformular³¹⁷.

Fehler melden: "We ask that you notify us immediately of any errors or incomplete entries that you discover, and we will make every effort to correct them." [Mar Jj].

K4.4 Beratung zum Forschungsdatenmanagement

Beratung: Kein Hinweis auf Webseite.

Hilfestellung: Kein Hinweis auf Webseite.

Verweis auf aktuelle Richtlinien / Standards: Kein Hinweis auf Webseite.

K5.1 Organisationsform

Passt zu Zielen / Aufgabe: Gehostet im Lamont-Doherty Earth Observatory der Columbia University [Lam J], fachliche Anbindung an Institut.

Langfristigkeit: Keine Hinweise auf Webseite.

Träger bekannt: "Funding for MGDS is provided by the National Science Foundation as part of IEDA Cooperative Agreement." [Mar Ja]. Hinweis im Footer: "Funded by the US National Science Foundation (NSF)".

Gemeinnützig: Keine Hinweise auf Webseite.

K5.2 Personalausstattung und -qualifizierung

9 Mitarbeiterx, 2 Projektleiterx [Lam J], alles Fachwissenschaftlerx, keine Informationswissenschaftlerx.

K5.3 Forschungs- und Entwicklungstätigkeiten

Steuert Weiterentwicklung der Infrastruktur: "The MGDS Relational database backend was designed as a fully integrated system to handle a suite of marine geoscience data while also serving the specific needs of the individual science communities that it services." [Mar Jf].

Kooperationen mit FDM- Initiativen und Fachdisziplin: Keine Hinweise auf Webseite.

Mitarbeit an FDM-Standards: Keine Hinweise auf Webseite.

Forschungsprojekte: Keine Hinweise auf Webseite.

K5.4 Technische Infrastruktur

Angemessen, Zuverlässig, Sicher, Skaliert: "The MGDS Relational database backend was designed as a fully integrated system to handle a suite of marine geoscience data while also serving the specific needs of the individual science communities that it services." [Mar Jf].

schnelle Lade- & Reaktionszeiten: "Most searches take just a few seconds, however, searches based on location may take longer." [Mar Ji].

Ausfallsicherheit (verteilte DB, Server): Keine Hinweise auf Webseite.

Back-Up Strategie: Keine Hinweise auf Webseite.

Überwachung der Komponenten: Keine Hinweise auf Webseite.

techn. Infrastrukt. dokumentiert: Keine Hinweise auf Webseite.

offene Software: Keine Hinweise auf Webseite.

Forschungseth. und rechtl. Richtlinien: Keine Hinweise auf Webseite.

K6.1 Gebühren für die Datenabgabe und -nutzung

Bekannte Gebührenmodelle (Dauer / Volumen Staffelung): Keine Hinweise auf Webseite.

Verträge zwischen FDR und Institut: Keine Hinweise auf Webseite.

K6.2 Finanzierung des Betriebs

Langfristiges Konzept: Keine Hinweise auf Webseite.

Budgetplanung: Keine Hinweise auf Webseite.

Prognosen für Zukunft: Keine Hinweise auf Webseite.

³¹⁷ <http://www.marine-geo.org/about/contact.php>

DR-NTU Data Repository

Beispieldatensatz für Screenshots:

Beispiel: <http://dx.doi.org/10.21979/N9/IP6DBD>

re3data Datensatz für die MGDL [re317a].

K1.1 Datenabgabe

Allgemein: "There is no file size limit to the entire dataset. However, depending on internet and network conditions, each single file upload is limited to 3.5 GB and a double zipped folder is limited to 1 GB (via LAN connection to NTU network)." [NTU17h].

Zulässiges Format: "DR-NTU (Data) will accept any file formats and no charges are imposed on the data depositor." [NTU17h].

Möglichkeiten des Uploads: Formular, "1. Login to DR-NTU (Data) using your NTU Staff Authentication and password. 2. Go to "Add Data"." [NTU17k]. Oder über Tabular Ingest (erlaubt SPSS, TATA, R, Excel XLSX, CSV): "Dataverse stores the raw data content extracted from such files in plain text, TAB-delimited files." [NTU17m].

Daten-Integritätsprüfung: "DR-NTU (Data) does not review all User Uploads before they are made available on the Site, or before they are Published." [NTU17i].

K1.2 Metadaten

Metadatenkategorien: Unterteilt in Citation, Domain specific und File-level Metadata [NTU17l].

standardisierte Formate: Diverse, volle Liste siehe Supported Metadata and References³¹⁸, für Citation Metadata u.a.: DDI Lite, DataCite 3.1, Dublin DCMI oder für Domäne Astrophysik das VOResource Schema. "Currently supported export formats are DDI, Dublin Core and JSON." [NTU17l]. "To aid discovery and reuse, DR-NTU (Data) provides about 100 metadata field choices for the depositor to describe the dataset. The metadata schemas used in Dataverse are compliant with international standards e.g. DDI Lite, DataCite 3.1, DDI 2.5 Codebook, Dublin Core's DCMI metadata terms, etc." [NTU17c].

unterstützende Tools: Kein Hinweis auf Webseite.

formularbasierte Eingabe: Ja, siehe How to upload Guideline³¹⁹ auf Seite 2. Hinweis auf Pflichtangaben: Markiert mit Sternchen.

Hilfe bei Schlagworten (Systematik): "Language field uses ISO 639-2 controlled vocabulary. [...] Country / Nation field uses ISO 3166-1 controlled vocabulary. [...] along with controlled vocabulary from subsets of the OBI Ontology and the NCBI Taxonomy for Organisms." [NTU17l].

1.3 Dokumentation

Dokumentation der FD: "[...] and each dataset contains descriptive metadata and data files (including documentation and code that accompany the data)." [NTU17a].

Verlinkung auf Datenpublikation: Keine Hinweise auf Webseite.

K1.4 Zitierfähigkeit

Richtlinien zur Zitation von FD: Zitationsvorschlag auf Datensatzebene. Richtet sich nach den Dataverse Community Norms: "Any materials [...] created that employ, reference, or otherwise utilize the data (in whole or in part) gathered from deposited datasets should credit the source with the applicable data citation generated by Dataverse (found on the dataset page). These citations include the data authors, data identifier, and other information in accordance with the Joint Declaration of Data Citation Principles (<http://force11.org/datacitation>) for all research

³¹⁸https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/DRNTUData_SupportedMetadataReferences-2e09yn2.pdf

³¹⁹https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/DRNTUData_HowToUpload-2bmgpd1.pdf

data." [NTU17e].

PID: "Each dataset will also be assigned a Digital Object Identifier (DOI). Hence, it is easier for people to cite your data and for you to share your data with others." [NTU17h].

K1.5 Lizenzmodelle

Nutzungsbedingungen festlegen: "Downloaders must be registered Users of the Site or guests that agree to the CC-BY-NC or the Custom Data Usage Agreement for a dataset in order to take advantage of the Site's Services, including downloading any materials or datasets." [NTU17i] (S. 7). "You acknowledge that DR-NTU (Data)'s default data usage license agreement for all uploaded materials is a CC-BY-NC. [...] Users also have the option of drafting a custom data usage license agreement. [...] Users also have the option of choosing to use DR-NTU (Data)'s restricted data usage license agreement ("Data Use Agreement")." [NTU17i] (S. 6f.).

offene Lizenzen anbieten: "The data usage license in DR-NTU (Data) is by default a Creative Common CC-BY-NC license. However, if you find the default license is not suitable due to some circumstances, e.g. agreements with external collaborators and parties, you have the option to use other Creative Common licenses or to draft a custom data usage license." [NTU17h].

Datenschutz: Verweis auf Data Protection and Privacy Statement der Universität³²⁰.

MD Bindungslos (CC0): Keine Hinweise auf Webseite.

K1.6 Verweise auf Veröffentlichungen

Link zu Veröffentlichung der Publikation: "A dataset record provides metadata, information about related publications and links to data files." [NTU17d]. Siehe Screenshot NTU002_Verlinkung.

Link zu Publikationen die nachnutzen: Keine Hinweise auf Webseite.

Verknüpfung zugehöriger Projekte: In Dataverses möglich³²¹.

K1.7 Nutzungsstatistik

Metriken auf Datensatzebene: Metriken über Google Analytic und Piwik implementiert, in Planung: Altmetrics [NTU17g]. Siehe Screenshot NTU004_Metriken.

Standardisierte Erhebung: Keine Hinweise auf Webseite.

K1.8 Anerkennung der wissenschaftlichen Leistung

Anreize für Forscherx: "There is no lack of literature that tells us about the positive correlation between publicly available research data and increased literature impact." [NTU Ja].

Intregation PID: Orcid in Planung [NTU17g].

Zitationsraten / Altmetrics: In Planung: Altmetrics [NTU17g] automatische Reportings: "The Site collects and stores Download Data from each download for all users (both registered and guest), which can then be downloaded and accessed by the Depositor." [NTU17i].

K2.1 Zugangsformen

öffentlich / Nutzergruppen: "If you restrict any files in your dataset, you will be prompted by a pop-up to enter Terms of Access for the data. This can also be edited in the Terms tab or selecting Terms in the "Edit" dropdown button in the dataset." [NTU17f] (S. 3).

Nutzerkreis Abgabe: "NTU researchers, staff and students who has a valid NTU network account can deposit and publish their research data at DR-NTU (Data)." [NTU17h].

Zugriffsteuerung (rechtliche / ethische Einschränkungen): s.o. NTU network account.

Metadaten restriktionsfrei: Kein Hinweis auf Webseite.

K2.2 Embargofunktion

Embargofunktion: Kein Hinweis auf Webseite.

³²⁰ <http://www.ntu.edu.sg/Pages/Privacy.aspx>

³²¹ <http://blogs.ntu.edu.sg/lib-datamanagement/about-dr-ntu-data/>

Embargofristen: Kein Hinweis auf Webseite.

Begründung & Ende sichtbar: Kein Hinweis auf Webseite.

K2.3 Recherchemöglichkeiten

Allgemein: Erlaubt Browsing. "In DR-NTU (Data), browsing happens when a user hasn't entered a search term or query into the basic search bar." [NTU17j].

Freie Textsuche: Einfache und Erweiterte Suche. "From the DR-NTU (Data) homepage, you can begin searching with an exact phrase search or entering a search term or query in the search box that says, 'Search this dataverse'. [...] To perform an advanced search, click the "Advanced Search" link next to the search bar. There you will have the ability to enter search terms for Dataverses, Dataset Metadata (citation and domain-specific), and file-level metadata." [NTU17j].

Suchoperatoren: Nicht enthalten in Suchhilfe. Test mit BOOLSchen Operatoren, Trunkierung und exakter Suche erfolgreich, Wildcard nicht erfolgreich (siehe Screenshot NTU006_Wildcard).

attributspezifisch (Metadaten): "In an advanced search, you can refine your criteria by choosing which Metadata fields to search. You can perform an advanced search on Citation metadata fields as well as domain specific metadata fields [...]" [NTU17j].

Facettierung: "Facets: to the left of the search results, there are several facets a user can click on to narrow the number of results displayed." [NTU17j].

Empfehlungssysteme: Kein Hinweis auf Webseite.

MD-Suche / Inhaltliche Suche: Kein Hinweis auf inhaltliche / Volltextsuche.

Indexierbare Landingpage: Im Kontext der Begriffsklärung genannt: "Publish – A function available for User Submissions by which Users can make their dataverse and/or their dataset containing data files publicly available and publicly searchable on the DRNTU (Data) search engine and third-party search engines (e.g., Google Search or Bing Search)." [NTU17i] (S. 1).

2.4 Integration in Zitationsdatenbanken

Metadaten bereit / liefert an Zit.-DB: Keine Hinweise auf Webseite.

MD-Abfrage-Schnittstellen (OAI-PMH): "OAI-PMH" [NTU17g]. OAI-PMH Schnittstelle existiert in DataVerse Software, sonst keine weiteren Hinweise gefunden.

K2.5 Programmierschnittstellen

APIs: "OAI-PMH, SWORD API (deposit), Search API, Data Access API (download), Native API (everything else)" [NTU17g]. Siehe oben, keine Informationen ob implementiert oder nur möglich.

K3.1 Policies

Richtlinien: "DR-NTU (Data) Policies³²², DR-NTU (Data) General Terms of Use³²³, allgemeine NTU Research Data Policy³²⁴. Außerdem allgemeines Data Protection and Privacy Statement der NTU³²⁵ und Sample Data Usage Agreement³²⁶.

K3.2 Datenvalidierung

"DR-NTU (Data) does not review all User Uploads before they are made available on the Site, or before they are Published." [NTU17i].

K3.3 Versionierung

³²²<http://blogs.ntu.edu.sg/lib-datamanagement/dr-ntu-data-policies/>

³²³https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/12/DRNTUData_GeneralTermsOfUse_v2_01122017-2azwwrc.pdf

³²⁴<http://research.ntu.edu.sg/riero/RI/Pages/Research-Data-Policies.aspx>

³²⁵<http://www.ntu.edu.sg/Pages/Privacy.aspx>

³²⁶<https://blogs.ntu.edu.sg/lib-datamanagement/files/2017/03/Sample-Data-Usage-Agreement%2D17m4cli.pdf>

Versionierung: "Versioning control: - Changes are tracked and recorded" [NTU17d]. "Once you have published a dataset, any metadata or file changes (e.g. by uploading a new file, changing file metadata, adding or editing metadata) will be tracked in our versioning feature." [NTU17f] (S. 9).

aktuelle V. gekennzeichnet / ältere V. verknüpft: Keine Hinweise auf Webseite.

Änderungen protokolliert: "To view what has exactly changed starting from the originally published version to any subsequent published versions: click on the Versions tab on the dataset page to see all versions and changes made for that particular dataset." [NTU17f] (S. 10). Siehe Screenshot NTU003_Versionierung.

PID-Vergabe deutlich: Keine Hinweise auf Webseite.

K3.4 Löschung

Geregelt: "Delete Dataverse: you are able to delete your dataverse as long as it is not published and does not have any draft datasets." [NTU17f] (S. 1). Alternativ: "Deaccessioning a dataset or a version of a dataset is a very serious action that should only occur if there is a legal or valid reason for the dataset to no longer be accessible to the public." [NTU17f] (S. 10).

URL & PID funktionsfähig: "A tombstone landing page with the basic citation metadata will always be accessible to the public if they use the persistent URL (Handle or DOI) provided in the citation for that dataset. Users will not be able to see any of the files or additional metadata that were previously available prior to deaccession." [NTU17f] (S. 10).

Grund der Löschung auf Landing-Page: Kein Hinweis auf Webseite.

K3.5 Sicherstellung langfristiger Zugänglichkeit

Übernimmt Verantwortung: "All data deposited in DR-NTU (Data) will be preserved and archived and there will be no charge to deposit your data into DR-NTU (Data)." [NTU17h].

Sicherungsmaßnahmen: "This includes, without restriction, permission to: [...] 4. store, translate, copy or re-format the Content in any way to ensure its future preservation and accessibility, and improve usability and/or protect respondent confidentiality; [...]" [NTU17i].

Sicherungsdauer: NTU Research Data Policy verpflichtet zu Aufbewahrung für 10 Jahre: "All research data related to a research project shall be retained for 10 years after publication or after the completion of the project, whichever is later." [Nan15]. Keine Aussage zur Sicherungsdauer im FDR

Nachfolgeregelungen: Kein Hinweis auf Webseite.

K3.6 Zertifizierung

Nicht zertifiziert, FDR wurde erst im September 2017 eröffnet [Wen17].

Einhaltung von Anforderungen: Nicht zutreffend.

Prüfberichte Online: Nicht zutreffend.

K4.1 Usability

Nutzerxfreundlichkeit: Keine Besonderheiten.

intuitive Bedienung: Ja.

etablierte Darstellungsformen: Ja.

flache Lernkurve: Eher ja, aber fehlende Suchhilfe. Barrierefreiheit: Browsererweiterung aXe findet 65 Verstöße gegen Barrierefreiheit (siehe Screenshot NTU007_Accessibility-Suche), vor allem fehlender Farbkontrast und fehlende Beschriftung von Links. 22 Verstöße auf Datensatzebene (siehe Screenshot NTU008_Accessibility-Datensatz), wieder fehlender Farbkontrast. Manuelle Betrachtung: Inhaltliche Beschreibung von Videos nicht vorgesehen, kann per Freitext in Description umgesetzt werden (Praxisbeispiel: Forscherx legt Textdatei mit Transkript bei, siehe Screenshot NTU009_BeispielTranskript). Keine Funktion für das Einstellen von Transkript oder Untertitel gefunden. Keine Textalternativen für Video, keine eigene Suche nach

Bildern oder Tondokumenten. Keine Anpassung Textgröße oder Farbschema. Nicht mit Tastatur allein bedienbar. Keine zeitlichen Einschränkungen. Keine Blitze. Keine Bemühungen hinsichtlich Lesbarkeit. Keine Eingabehilfe, keine Fehlerkorrektur.

K4.2 Darstellung der Datensätze

Datenvisualisierung: "GeoConnect is the platform integrating Dataverse and WorldMap together and what you will use to visualize your data." ([NTU17b, S. 3]), sowie TwoRavens für tabellarische Daten (ebd. S. 1). Dazu jedoch keine Funktion auf Webseite gefunden.

Vorschaufunktion: Keine Vorschaufunktion für Video.

K4.3 Support zu Repositoriumsaktionen

Hinweis: Die Hauptseite der Hilfedokumente³²⁷ gibt an, dass die User Guides von den Dataverse Guides adaptiert wären. Es scheint aber in weiten Teilen direkter Wortlaut des Dataverse Projektes³²⁸ ohne korrekte Quellenangabe zu sein. In den meisten Fällen wurden durch die NTU-Mitarbeiterx die Dateinamen und einige Überschriften geändert. Die Dateien werden in dieser Arbeit dennoch mit den NTU-Angaben und Verlinkungen zitiert wie sie auf der Webseite der Universität erscheinen. Betroffen sind die Quellen [NTU17f], [NTU17i], [NTU17j], [NTU17b] und [NTU17m].

Hilfeseiten: Upload Guide[NTU17k], User Guides zu diversen Themen³²⁹ (siehe Hinweis), FAQ³³⁰.

Schulungsangebote: Angekündigt: "Apart from the regular DMP workshops, we would be lining up some workshops to highlight the features and let participants have a hand-on experience of using DR-NTU (Data)." [Wen17].

Supportanfragen: Über Kontaktformular³³¹.

Fehler melden: Keine Hinweise auf Webseite.

K4.4 Beratung zum Forschungsdatenmanagement

Beratung: "6.4 LIBRARY 6.4.1 Provide training and advice on research data management including metadata description, DMP preparation, copyright and licenses." [Nan15]. Angebote: Einführung in FDM, FD Definition, DMP Erklärung und Workshops usw.³³². "These [Office of Information, Knowledge & Library Services data curation] services include advice, review and training on metadata, file organisation, data documentation, terms of use open formats and open access." [Aso17].

Hilfestellung: "We understand that first-time users may have many queries on the data deposit process. Fear not, we do provide a mediated deposit service. All you have to do is to contact the friendly research data librarians at NTU Libraries for help!" [Wen17]. Vorlage für DMP³³³.

Verweis auf aktuelle Richtlinien / Standards: Verweisen auf diverse Standards, z.B. RCUK Common Principles on Data Policy [NTU Ja]. und die Harvard Dataverse Policies³³⁴. Verweis auf andere FDR, re3data, Datenjournale [NTU17n].

K5.1 Organisationsform

Passt zu Zielen / Aufgabe: Angeschlossen an Universität, zuständig ist Bibliothek, betreut von IT-Abteilung der Universität [Nan15].

Langfristigkeit: Keine Hinweise auf Webseite.

³²⁷ <http://blogs.ntu.edu.sg/lib-datamanagement/dr-ntu-data-user-guides/>

³²⁸ <http://guides.dataverse.org/en/4.3/user/>

³²⁹ <http://blogs.ntu.edu.sg/lib-datamanagement/dr-ntu-data-user-guides/>

³³⁰ <http://blogs.ntu.edu.sg/lib-datamanagement/dr-ntu-data-faq/>

³³¹ <https://blogs.ntu.edu.sg/lib-datamanagement/contact-us/>

³³² <https://blogs.ntu.edu.sg/lib-datamanagement>

³³³ <http://blogs.ntu.edu.sg/lib-datamanagement/dmp/ntu-data-management-plan>

³³⁴ <https://dataverse.org/best-practices/harvard-dataverse-policies>

Träger bekannt: Angeschlossen an Universität, zuständig ist Bibliothek, betreut von IT Abteilung der Universität [Nan15]. "DR-NTU (Data) [...] is the institutional open access research data repository for Nanyang Technological University (NTU) supported by the Library and CITS." [NTU17a].

Gemeinnützig: Keine Hinweise auf Webseite.

K5.2 Personalausstattung und -qualifizierung

"The Research Data Management Team is part of the Office of Information, Knowledge & Library Services." [NTU Jb], dort 3 Mitarbeiterx angezeigt, keine Informationen zur Qualifizierung.

K5.3 Forschungs- und Entwicklungstätigkeiten

Steuert Weiterentwicklung der Infrastruktur: Keine Hinweise auf Webseite.

Kooperationen mit FDM- Initiativen und Fachdisziplin: Keine Hinweise auf Webseite.

Mitarbeit an FDM-Standards: Keine Hinweise auf Webseite.

Forschungsprojekte: Keine Hinweise auf Webseite.

K5.4 Technische Infrastruktur

Angemessen, Zuverlässig, Sicher, Skaliert: "It is a protected and secure repository that allows you to deposit your final research data and to make your data more visible and discoverable." [NTU17h].

schnelle Lade- & Reaktionszeiten: Keine Hinweise auf Webseite.

Ausfallsicherheit (verteilte DB, Server): Keine Hinweise auf Webseite.

Back-Up Strategie: Keine Hinweise auf Webseite.

Überwachung der Komponenten: Keine Hinweise auf Webseite.

techn. Infrastrukt. dokumentiert: Keine Hinweise auf Webseite.

offene Software: "[...] built on an open-source web application called Dataverse which is developed and used by Harvard University." [NTU17a].

Forschungseth. und rechtl. Richtlinien: Keine Hinweise auf Webseite.

K6.1 Gebühren für die Datenabgabe und -nutzung

Bekannte Gebührenmodelle: "All data deposited in DR-NTU (Data) will be preserved and archived and there will be no charge to deposit your data into DR-NTU (Data)." [NTU17h].

Verträge zwischen FDR und Institut: Keine Hinweise auf Webseite.

K6.2 Finanzierung des Betriebs

Langfristiges Konzept: Keine Hinweise auf Webseite.

Budgetplanung: Keine Hinweise auf Webseite.

Prognosen für Zukunft: Keine Hinweise auf Webseite.

A.3. Vergleichstabellen

Die Quellen der enthaltenen Angaben befinden sich in Anhang A.2.

K1 Veröffentlichung und Nachnutzung

Kategorie / FDR	Stanford Digital Repository	Phaidra	Zenodo	Brown Digital Repository	LSE Research Online	OceanRep GEOMAR Repository	Marine-Geo Digital Library	DR-NTU Data Repository
K1.1 Datenabgabe Zulässiges Format	Alle	Alle	Alle	Alle	n.a.	Explizit nur: PDF, Postscript, HTML, und ASCII	Liste von Formaten	Alle
Uploadmöglichkeiten	Browser, Bibliothekarx	Browser, Importer	Browser, Schnittstelle	n.a.	Bibliothekarx, Forscherx Methode unbekannt	Bibliothekarx, Browser, Import	Browser, Google Form u. Filesharing	Browser
Daten-Integritätsprüfung	n.a.	n.a.	Nein	Bibliothekarx	Bibliothekarx	Bibliothekarx	n.a.	Nein
K1.2 Metadaten Struktur der Metadaten	n.a.	Beschreibend, strukturell, technisch, administrativ	n.a.	Beschreibend, Administrativ, Technisch, Strukturell, Aufbewahrung, Rechte	n.a.	n.a.	Grundlegend, örtlich und zeitlich, Verlinkung, Expeditionen	Bibliografisch, Domänenspezifisch, Dateiebene
Standardisiertes Format	MODS, DC	DC, LOM, METS	JSON Schemata, DataCite-konform	DC, MODS, METS u.a.	n.a., Export in DC, METS etc.	n.a.	ISO19115	Div., u.a. DDI, DC, JSON
Unterstützende Tools	Empfehlungen	Ingest- Applikation, Vorlage	n.a.	n.a.	In Planung: Autovervollständigung MD Felder	Autovervollständigung, Vorlage	Autovervollständigung	n.a.
Formularbasiert	Ja	Ja	Ja	n.a.	n.a.	Ja	Ja	Ja
Hinweis auf Pflichtangaben	Ja	Ja	Ja	n.a.	n.a.	Ja	Ja	Ja
Hilfe bei Schlagworten	Nein	Thesauri, Klassifikationen	Nein, andere Felder mit Vokabular	n.a.	Kontrolliertes Vokabular	Kontrolliertes Vokabular	n.a.	Einzelne Felder mit ISO, Ontologien
K1.3 Dokumentation FD-Dokumentation Link zu Datenpublikation	Verknüpfung o. Datei Möglich	In Entwicklung n.a.	Datei Ja	Datei Möglich	n.a. n.a.	n.a. Möglich	Möglich Ja	Datei n.a.
K1.4 Zitierfähigkeit Zitationsrichtlinie	nur Zitationsvorschlag	n.a.	Data Citation Principles, Zitationsvorschlag DOI	nur Zitationsvorschlag DOI	nur Zitationsvorschlag DOI	nur Zitationsvorschlag DOI, URN	Zitationsrichtlinie, Zitationsvorschlag DOI	Data Citations Principles, Zitationsvorschlag DOI
Persistent ID	PURL mit DRUID	Handle, DOI	DOI	DOI	DOI	DOI, URN	DOI	DOI
K1.5 Lizenzmodelle Nutzungsbedingungen offene Lizenzen Datenschutz	Lizenz angezeigt CC u. ODC n.a.	Terms of Use, Policy CC Nutzungsbedingungen	Policy CC Terms of Use	Policy CC, GNU GPL n.a.	FAQ CC Terms of Use des Trägers	Suchhilfe CC n.a.	Terms of Use CC n.a.	Terms of Use CC, Maßgeschneidert Data Protection and Privacy Statement des Trägers n.a.
MD Bindungslos	n.a.	Ja	Ja	n.a.	n.a.	Nein	Nein	n.a.
K1.6 Verweise auf Veröffentlichungen Link zu Publikation Link zu nachnutzenden Pub. Verknüpfung mit Projekt / FD	Ja Möglich Related Content	Ja n.a. Ja, Unspezifiziert	Ja Möglich Sammlungen, Communities	Möglich Möglich Sammlungen	Ja n.a. Sammlungen	Ja n.a. Forschungsbereich u. autom. durch FDR	Ja n.a. Ja, Unspezifiziert	Ja n.a. Dataverses
K1.7 Nutzungsstatistik Wird Erhoben	Nein	Ja	n.a.	In Planung	Ja	Ja	Ja, nicht öffentlich	Ja

Fortsetzung auf der nächsten Seite ...

...Fortsetzung von K1 Veröffentlichung und Nachnutzung

Kategorie / FDR	Stanford Digital Repository	Phaidra	Zenodo	Brown Digital Repository	LSE Research Online	OceanRep GEOMAR Repository	Marine-Geo Digital Library	DR-NTU Data Repository
Anzeige Datensebene	n.a.	Ja	n.a.	n.a.	Ja	Ja (+ Gesamtstatistik)	Nein (nur Gesamtstatistik)	Ja
standardisierte Erhebung	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	Google Analytics, Piwik
K1.8 Anerkennung der wissenschaftlichen Leistung Anreize für Forscherx	Sichtbarmachung	n.a.	Sichtbarmachung, Anerkennung	Sichtbarmachung, Anerkennung	Sichtbarmachung, Vernetzung, Aktualität (RSS), Publikationsliste	Sichtbarmachung, Anerkennung, Publikationsliste	n.a.	Sichtbarmachung
Intregation PID	Nein	n.a.	Ja	n.a.	Nein	Ja (+ GND, ResearcherID, Scopus ID)	n.a.	In Planung
Zitationsraten / Altmetrics	Altmetrics in Planung	n.a.	Altmetrics	Verweis auf externe Angebote	Altmetrics	n.a.	Nicht öffentlich, nur für Autorx	Altmetrics in Planung
Reporting an Förderer	n.a.	n.a.	Europäische Kommission	n.a.	n.a.	Interne Förderung, Jahresberichte, Evaluierung	halbjährlich Zugriffs-/Downloadzahlen an Forscherx	Downloadzahlen für Forscherx

Tabelle A.5.: Vergleich K1 Veröffentlichung und Nachnutzung

K2 Zugriff und Auffindbarkeit

Kategorie / FDR	Stanford Digital Repository	Phaidra	Zenodo	Brown Digital Repository	LSE Research Online	OceanRep GEOMAR Repository	Marine-Geo Digital Library	DR-NTU Data Repository
K2.1 Zugangsformen Beschränkung	Öffentlich, Institutsangehörige	Öffentlich, Eingeschränkt	Öffentlich, Eingeschränkt, Geschlossen	Öffentlich, Institutsangehörige, Geschlossen, Eingeschränkt	n.a.	Öffentlich, Institutsangehörige, Eingeschränkt, Geschlossen	Öffentlich, Geschlossen	Öffentlich, Eingeschränkt
Nutzerkreis Abgabe	Institutionsangehörige	Institutionsangehörige, auch externe Projektmitarbeiter	Alle	Institutionsangehörige, Studierende nur in Projekten	Institutionsangehörige ab Doktorandx	Institutionsangehörige	Alle, hauptsächlich Expeditionsteilnehmerx	Institutionsangehörige
Zugriffsteuerung	n.a.	Nutzergruppe	Nutzergruppe, begrenzter Zugriff	Nutzergruppe	n.a.	Begrenzter Zugriff	Passwort	n.a.
Metadaten immer sichtbar	n.a.	Ja	Ja	n.a.	n.a.	Ja	n.a.	n.a.
K2.2 Embargofunktion Embargo möglich	Ja	Ja	Ja	technisch möglich, für FD n.a.	Ja	Ja	Ja	n.a.
Embargofrist Begründung & Ende sichtbar	3 Jahre n.a.	n.a. n.a.	Keine Begrenzung Ende sichtbar	n.a. n.a.	min. 2 Jahre n.a.	Keine Begrenzung n.a.	n.a. Ende sichtbar	n.a. n.a.
K2.3 Recherchemöglichkeiten Freie Textsuche	Einfach u. Erweitert	Einfach u. Erweitert	Einfach, Schnittstelle in Planung	Einfach, Schnittstelle	Einfach u. Erweitert	Einfach, Erweitert, Schnellsuche	nur externer Metabrowser, Schnittstelle n.a.	Einfach u. Erweitert, Schnittstelle
Suchoperatoren	BOOL, Trunkierung, Wildcard	Test: BOOL, Trunkierung, Wildcard erfolgreich	Test: BOOL erfolgreich, Trunkierung, Wildcard ohne Fehlermeldung	Test: BOOL, Trunkierung, Wildcard ohne Fehlermeldung	Test: BOOL erfolgreich, Trunkierung, Wildcard ohne Fehlermeldung	BOOL, Test: Trunkierung ohne Fehlermeldung, Wildcard nicht möglich	n.a.	Test: BOOL erfolgreich, Trunkierung ohne Fehlermeldung, Wildcard nicht möglich
Attributspezifische Suche	Ja	Ja	Ja	Auswahl	Auswahl	Ja	Einfach, Erweitert, Karte	Ja
Facettierung	Ja	Ja	Ja	Wenige	Nein	Nein	Nur extern	Ja
Empfehlungssysteme	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Fortsetzung auf der nächsten Seite ...

...Fortsetzung von K2 Zugriff und Auffindbarkeit

Kategorie / FDR	Stanford Digital Repository	Phaidra	Zenodo	Brown Digital Repository	LSE Research Online	OceanRep GEOMAR Repository	Marine-Geo Digital Library	DR-NTU Data Repository
Volltextsuche Indexierbare Landingpage	Nein Ja	Nein Ja	Nein Known-Item Suche erfolgreich	In Planung Nur Öffentliche	Möglich Ja	Ja Ja	Nein Known-Item Suche nicht erfolgreich	Nein Ja
K2.4 Integration in Zitationsdatenbanken Metadaten bereit für Zit.-DB Metadaten-Schnittstelle	n.a. n.a.	n.a. OAI-PMH, Base URL separat	DataCite u.a. OAI-PMH, REST	n.a. n.a.	n.a. OAI-PMH	Ja, un spezifiziert OAI-PMH	n.a. MGDS FileServer	n.a. OAI-PMH, keine Base URL
K2.5 Programmierschnittstellen APIs Funktion	RESTful, IIIF Zugang, un spezifiziert	phaidra-api (REST), OAI-PMH. LTI in Planung Ingest, MD u. FD Harvesting	OAI-PMH, REST Ingest, MD Harvesting (Datensatz u. Sammlung). In Planung: FD Harvesting, Suche	Unspezifiziert Suche, MD Harvesting (Datensatz u. Sammlung)	OAI-PMH MD Harvesting	OAI-PMH MD Harvesting	REST, OGC MD u. FD Harvesting, Suche	potenziell OAI-PMH, SWORD, Eigenentwicklungen möglich Diverse möglich, Umsetzung unbekannt

Tabelle A.6.: Vergleich K2 Zugriff und Auffindbarkeit

K3 Transparenz und Qualitätssicherung

Kategorie / FDR	Stanford Digital Repository	Phaidra	Zenodo	Brown Digital Repository	LSE Research Online	OceanRep GEOMAR Repository	Marine-Geo Digital Library	DR-NTU Data Repository
K3.1 Policies Richtlinien	Deposit R. (+ Hilfen Seiten)	Allgemeine R. (+ Terms of Use, Träger OA Policy)	Allgemeine R. (+ Terms of Use, FAIR Principles)	Allgemeine R.	Mission Statement in FAQ, Takedown R. (Träger OA + FDM R. + Terms of Use)	Allgemeine R.	nur Terms of Use, Submission Guideline	Allgemeine R. (+ Terms of Use, Träger Policies)
K3.2 Datenvalidierung Vollständigkeit der MD Verantwortung	n.a. n.a.	n.a. n.a.	Nein Datengeberx	Ja n.a.	Ja n.a.	Ja Datengeberx	n.a. n.a.	Nein Datengeberx
K3.3 Versionierung Versionierung möglich Anzeige u. Verknüpfung Änderungsprotokoll PID-Vergabe	Ja Nur aktuelle V. Online Ja Nur aktuelle V. Online	Ja Versionsanzeige, alle V. verlinkt n.a. Alte V. behält PID	Ja Versionsanzeige, übergeordneter PID n.a. Ein PID pro V. u. ein Übergeordneter	Ja n.a. Ja Alle V. unter einem PID	n.a. n.a. n.a. n.a.	Ja alle V. verlinkt n.a. PID resolved auf aktuelle V.	n.a. n.a. n.a.	Ja Versionsanzeige Ja n.a.
K3.4 Löschung Geregelt URL & PID bleiben aktiv Grund der Löschung angezeigt	Nein n.a. n.a.	Ja n.a. n.a.	Ja Ja Ja	Ja n.a. n.a.	Nein Möglich n.a.	Ja Ja Möglich	Nein n.a. n.a.	Ja Ja n.a.
K3.5 Sicherstellung langfristiger Zugänglichkeit Übernimmt Verantwortung Sicherungsmaßnahmen Sicherungsdauer	Ja Redundant, Ortsverteilt, Bandsicherung, Prüfung n.a.	Ja Verweis auf OAIS, Migration Datenträger und Format min. 5 Jahre	Ja Ortsverteilt, redundant, tägliche Bandsicherung, Prüfung min. 20 Jahre	Ja Ortsverteilt, Migration, Prüfung n.a.	Ja Elektronische Übersetzung, Redundant n.a.	Ja Backups, un spezifizierte technische Änderungen Unbestimmt	Ja n.a. n.a.	Ja Elektronische Übersetzung, Formatmigration, Redundant n.a., möglich 10 Jahre

Fortsetzung auf der nächsten Seite ...

...Fortsetzung von K3 Transparenz und Qualitätssicherung

Kategorie / FDR	Stanford Digital Repository	Phaidra	Zenodo	Brown Digital Repository	LSE Research Online	OceanRep GEOMAR Repository	Marine-Geo Digital Library	DR-NTU Data Repository
Nachfolgeregelung	n.a.	Zurück an Datengeberx	Unspezifizierte Repo., institutionell oder disziplinar	n.a.	n.a.	Unspezifiziertes angemessenes Repo.	n.a.	n.a.
K3.6 Zertifizierung Ist zertifiziert	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein

Tabelle A.7.: Vergleich K3 Transparenz und Qualitätssicherung

K4 Nutzererlebnis und Support

Kategorie / FDR	Stanford Digital Repository	Phaidra	Zenodo	Brown Digital Repository	LSE Research Online	OceanRep GEOMAR Repository	Marine-Geo Digital Library	DR-NTU Data Repository
K4.1 Usability Nutzererfreundlichkeit	n.a.	Mehrsprachig, MD Vorlagen	Unspezifische Fehlermeldung	n.a.	Datenimport nach ORCID, Auto. Aktualisierung Publikationslisten	Auto. Aktualisierung Publikationslisten, MD Vorlagen, Fehlermeldungen im Ingest	Alternative zur Kartensuche, Anzeige neuester Einträge	n.a.
Intuitive Bedienung	Ja	Ja	Ja	Ja	Nein	Nein	Karte: Nein, Text: Eher ja	Ja
Etablierte Darstellungsform Flache Lernkurve	Ja Ja, keine Tooltips	Ja Beeinträchtigt, keine Suchhilfe	Ja Beeinträchtigt, keine Suchhilfe	Ja Beeinträchtigt, keine Suchhilfe	Veraltet Beeinträchtigt, keine Tooltips	Veraltet Ja, Tooltips	Ja Beeinträchtigt, eingeschränkte Suchhilfe	Ja Beeinträchtigt, keine Suchhilfe
Barrierefreiheit	Nein	Nein, Suchvorschläge	Nein	Nein, Closed Captioning Service, Accessibility Prüfung geplant	Nein, Web Services Team als Ansprechpartner	Nein	Nein, Eingabehilfe	Nein
K4.2 Darstellung der Datensätze Datenvisualisierung	n.a.	einzelne Ressourcen (Buch)	n.a.	div. Bildanzeigen, Expertxberatung	n.a.	n.a.	Diverse geografische u. zeitliche, 3D Browser	Geografische u. Tabellarische Daten möglich
Vorschaufunktion	Bild u. Video	Bild u. Video	Nur Bild	Bild u. Video	Bild u. Video	Bild u. Video	Nur Bild	Nein
K4.3 Support zu Repositoriums-funktionen Hilfeseiten, FAQs	Diverse Hilfes. u. FAQ	Keine, nur Themenseiten	FAQ	FAQ, Informationsseiten	FAQ	FAQ, Hilfeseite, Suchhilfe	FAQ, Suchhilfe, Themenseiten	Dataverse Hilfsdokumente, FAQ, Uploadhilfe
Schulungsangebote Supportanfragen Fehler melden	Ja Kontaktformular Nein	Ja E-Mail Adresse Nein	Nein Kontaktformular Nein	Ja E-Mail Adresse Nein	Nein E-Mail Adresse Nein	Nein E-Mail Adresse Nein	Ja Kontaktformular Ja	Angekündigt Kontaktformular Nein
K4.4 Beratung zum Forschungsdaten- management Beratung Hilfestellung	Ja Ja	Ja Ja, Sandbox-Umgebung	Nein Nein	Ja Ja	Ja Ja	Ja Ja	Nein Nein	Ja Ja
Verweis auf Richtlinien / Standards	Ja	Ja	Nein	Ja	Ja	Ja	Nein	Ja

Tabelle A.8.: Vergleich K4 Nutzererlebnis und Support

K5 Organisation und Infrastruktur

Kategorie / FDR	Stanford Digital Repository	Phaidra	Zenodo	Brown Digital Repository	LSE Research Online	OceanRep GEOMAR Repository	Marine-Geo Digital Library	DR-NTU Data Repository
K5.1 Organisationsform Art	Abteilung in Bibliothek	Abteilung in Bibliothek, IT-Abteilung	IT-Abteilung	Abteilung in Bibliothek	Abteilung in Bibliothek	Abteilung in Bibliothek u. Datenmanagement Team	Forschungsinstitut	Abteilung in Bibliothek, IT-Abteilung
Langfristigkeit	n.a.	n.a.	Träger für >20 Jahre gefördert	n.a.	n.a.	n.a.	n.a.	n.a.
Träger bekannt	Universität	Universität	CERN, Europäische Kommission	Universität	Universität	Exzellenz Cluster, Sonderforschungsgebiete	National Science Foundation	Universität
Gemeinnützig	Ja	n.a.	n.a.	n.a.	n.a.	Träger: Ja	n.a.	n.a.
K5.2 Personalausstattung und -qualifizierung Anzahl MA Qualifizierung MA	30 Dienstleistung, LZA, IT	3 n.a.	15 n.a.	11 Informationswissenschaftlerx	4 Bibliothekrx, Library Assistent	8 n.a.	11 Fachwissenschaftlerx	>3 n.a.
K5.3 Forschungs- und Entwicklungstätigkeiten Infrastruktur	Entwicklung Rep. System	Entwicklung Rep. System, National u. International	International	Entwicklung Rep. System	n.a.	n.a.	Entwicklung Datenbank	n.a.
FDM- Initiativen und Fachdisziplin FDM-Standards	Dig. Sammlungen Manifesto for Linked Libraries	Vernetzung Leitungsebene n.a.	Div. EU Projekte Joint Declaration of Data Citation Principles	Digitalisierungsprojekte Gremien, Testumgebungen	n.a.	n.a.	n.a.	n.a.
Forschungsprojekte	Diverse	Infrastruktur, National	International, Fachspezifisch	Trendbeobachtung, Frameworks, Werkzeuge	n.a.	Framework, Vernetzung	n.a.	n.a.
K5.4 Technische Infrastruktur Allgemein	Hydra (Fedora-Basis), Moab	Phaidra (Fedora-Basis), solr, Linux OS	Invenio	Fedora 3, Linux OS	EPrints 3	EPrints	Relationale Datenbank	Dataverse
Lade- & Reaktionszeiten	Schnell	Schnell	Schnell, wird weiter optimiert	Schnell	Langsame Suche	Langsame Suche	Kartensuche langsam, Textsuche unauffällig	Schnell
Ausfallsicherheit	n.a.	n.a.	getrennte Server u. Systeme	getrennte Server	n.a.	n.a.	n.a.	n.a.
Back-Up Strategie	Bandsicherung, unspezifiziert	Nächtliche Bandsicherung für FD, redundant, Metadaten alle 12 Std.	n.a.	n.a.	n.a.	Regelmäßig, unspezifiziert	n.a.	n.a.
Überwachung der Komponenten techn. Infrastrukt. dokumentiert	n.a. n.a.	n.a. Github, für System u. API	Ja Github für System, Dokumentation der API	n.a. n.a.	n.a. n.a.	n.a. n.a.	n.a. n.a.	n.a. n.a.
offene Software	Ja	Ja	Ja	Ja	Ja	Ja	n.a.	Ja

Tabelle A.9.: Vergleich K5 Organisation und Infrastruktur

K6 Gebühren und Finanzierung

Kategorie / FDR	Stanford Digital Repository	Phaidra	Zenodo	Brown Digital Repository	LSE Research Online	OceanRep GEOMAR Repository	Marine-Geo Digital Library	DR-NTU Data Repository
K6.1 Gebühren für die Datenabgabe und -nutzung Gebührenmodell Vertrag zwischen FDR und Inst.	In Entwicklung n.a.	Kostenfrei n.a.	Kostenfrei, Spenden n.a.	n.a. n.a.	n.a. n.a.	Kostenfrei n.a.	n.a. n.a.	Kostenfrei n.a.
K6.2 Finanzierung des Betriebs Langfristiges Konzept	n.a.	n.a.	gebunden an CERN	2015-2017 enthalten im Technologieplan, bisher keine Aktualisierung n.a.	n.a.	Mitfinanziert von Exzellenz-Cluster und zwei Sonderforschungsbereichen ca. 3000€/Jahr in 2014 n.a.	n.a.	n.a.
Budgetplanung	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Prognose	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Tabelle A.10.: Vergleich K6 Gebühren und Finanzierung

A.4. Screenshots

Erste Analyse der FDR aus re3data

Vor allem referenziert in Anhang [A.1](#) und [A.2](#). Liegen der Arbeit im gleichnamigen Verzeichnis auf dem Datenträger bei. Die Benennung erfolgte nach dem Schema laufende Nummer_Name des FDR_optional: Art des Beleges. Beispiele: 002_GAMS, 042_DR-NTU-ErgebnisseVideo

Analyse der Kriterienkonformen FDR

Vor allem referenziert in Kapitel [5](#) und Anhang [A.1](#) und [A.2](#). Liegen der Arbeit im gleichnamigen Verzeichnis auf dem Datenträger bei. Die Benennung erfolgte nach dem Schema Abkürzung des FDR_Art des Beleges. Beispiele: BDR001_AnzeigeDatensatz, ORE011_Accessibility-Datensatz

Darin enthalten ist auch der Ordner "Push-Pull Prüfung".