

Fachrepositorium Lebenswissenschaften: Ausbau des Repositoriums zur Publikation von Forschungsdaten

Birte Lindstädt¹

ZB MED Leibniz-Informationszentrum Lebenswissenschaften

Das Fachrepositorium Lebenswissenschaften von ZB MED wird von einem „klassischen“ Repository für Graue Literatur und elektronische Zweitveröffentlichungen zu einem Aufbewahrungsort für eine Vielzahl unterschiedlicher Publikationsarten, u.a. Forschungsdaten, ausgebaut. Die Konzeption und die erforderlichen Metadaten zur Erfassung und Publikation von Forschungsdaten müssen die Spezifika dieser Publikationsart berücksichtigen. Der Beitrag stellt den Prozess der Weiterentwicklung des Fachrepositoriums zu einem Forschungsdatenrepositorium dar und zeigt die im Zuge dessen erfolgten Entscheidungen, Arbeitsschritte und Diskussionen auf.

The ZB MED Repository Life Science will be developed from a repository for grey literature and for articles that have already been published to a deposit for a variety of possibilities for publishing research results, e.g. research data. The concept and the necessary metadata for registration and publication of research data have to refer to the specification of this kind of publishing. The article shows the enhancements from a repository mainly for text publications to a research data repository and points out the decisions, steps and discussions connected with this process.

Einführung

Angebote zum Open-Access-Publizieren sind aus Sicht des ZB MED Leibniz-Informationszentrum Lebenswissenschaften ein wichtiger Baustein im Leistungsspektrum einer Zentralen Fachbibliothek, um den Gedanken des offenen Zugangs zu Wissen zu verbreiten und entsprechende Publikationen zu ermöglichen. ZB MED bietet daher für alle Publikationsarten vom Zeitschriftenartikel über Kongressbeiträge und Handbücher bis zu Forschungsdaten die Möglichkeit der Open-Access-Publikation auf „goldenen“ oder „grünen“ Weg. Diese Angebote stellen eine

¹ Robin Rothe als studentischer Mitarbeiter bei ZB MED hat im Wesentlichen die grundlegenden Recherche- und Analysetätigkeiten für den Ausbau des Fachrepositoriums Lebenswissenschaften erledigt und somit auch die Grundlagen für diesen Beitrag geliefert.

Alternative zu den kommerziellen Verlagsangeboten dar, um allen qualitativ geeigneten Forschungsergebnissen eine Publikationsmöglichkeit zu eröffnen. Alle Services sind unter dem Publikationsportal PUBLISSO zusammengefasst.

Das Fachrepositorium Lebenswissenschaften von ZB MED bietet Autorinnen und Autoren derzeit die Möglichkeit der kostenfreien, elektronischen Zweitveröffentlichung („grüner“ Weg) ihrer wissenschaftlichen Texte aus den Fachgebieten Medizin, Gesundheitswesen, Ernährungs-, Umwelt- und Agrarwissenschaften nach den Grundsätzen des Open Access. Der Schwerpunkt lag bisher auf Monographien und sog. Grauer Literatur.

Mit der Umstellung und Weiterentwicklung der Software durch den technischen Partner hbz (Hochschulbibliothekszentrum Nordrhein-Westfalen) ging 2015 auch eine konzeptionelle Neuorientierung einher, die die Erweiterung des Fachrepositoriums zur Aufnahme von unselbstständigen Veröffentlichungen in Form von Zeitschriftenartikeln, Kongressbeiträgen, Buchartikeln und Forschungsdaten beinhaltet. Darüber hinaus ist es künftig möglich, eine Filterung nach Institutionen vorzunehmen, so dass es auch als institutionelles Repositorium genutzt werden kann.

Für die neu hinzukommenden Publikationsarten wurde jeweils ein eigenes Metadatenschema erarbeitet. Eine besondere Aufgabe stellte dabei die Erarbeitung eines Schemas für Forschungsdaten dar, da hierbei andere Kriterien als bei reinen Texten eine Rolle spielen.

So muss ggf. eine Verknüpfung zu einer Textpublikation möglich sein, falls es sich nicht um singuläre Forschungsdaten handelt. Die Erklärung und damit Nutzbarkeit der Daten muss durch Erfassung beispielsweise von Abstract, Erhebungszeitraum oder -punkt, den Erfassungsort, ggf. Koordinaten etc. gewährleistet sein. Bei der Beschreibung der Daten selbst spielt das Format, die Größe und der Ressourcentyp eine Rolle.

Die Softwareumstellung auf das Fedora-System kombiniert mit dem Frontendsystem Drupal bietet die Möglichkeit verschiedenste digitale Materialien in unterschiedlichen Formaten zu speichern und wiederzugeben. Auch Datenmengen bis zu einem Gigabyte stellen für das System im Hinblick auf Forschungsdaten kein Problem dar. Wichtig für die Funktion als Forschungsdatenrepositorium ist darüber hinaus, dass das System Relationen digitaler Objekte untereinander abbilden kann.

So können auch Objekte aufgenommen werden, die sich aus verschiedenen Komponenten zusammensetzen und verschachtelte hierarchische Strukturen aufweisen. Die Objekte können mit erweiterten Metadaten, z.B. einem DOI oder Lizenzbestimmungen, versehen werden und Zugangsbeschränkungen erhalten. Diese Funktionen wie auch das integrierte Viewer-Modul sind elementare Grundlagen für die Integration von Forschungsdaten, um beispielsweise eine Beziehung zwischen einem Artikel und den zugehörigen Daten herzustellen oder Audio- und Videodateien wiederzugeben (Quast, A., 2016).

Fachrepositorien und Forschungsdaten

Das Fachrepositorium Lebenswissenschaften war und ist für ZB MED ein wichtiger Baustein im Rahmen des Bestandsaufbaus im Bereich digitaler Fachinformation. Das Fachrepositorium für Open-Access-Publikationen hat dabei zwei wesentliche Ausgangspunkte:

Zum einen dient es als Veröffentlichungsserver der Medizinischen Fakultät der Universität zu Köln, der es den Promovierenden und Habilitanden der Medizin ermöglicht, ihrer Veröffentlichungspflicht nachzukommen. Zum anderen bietet es bisher überwiegend Grauer (digitalisierter) Literatur aus den Forschungs- und Infrastruktureinrichtungen des lebenswissenschaftlichen ZB MED-Fächerspektrums einen zentralen Ort zur Einsicht und dauerhaften Auffindbarkeit.

Hinzu kommt die Funktion als Publikationsort für elektronische Zweitveröffentlichungen überwiegend zunächst als Print erschienener Literatur.

Während die Integration weiterer unselbstständiger Publikationsformen wie Zeitschriftenartikel oder Kongressbeiträge von anderen Fachrepositorien² oder auch institutionellen Repositorien angeboten wird, ist die Publikation von Forschungsdaten im Rahmen eines Open-Access Repositoriums noch nicht weit verbreitet. Beispiele finden sich hier bei institutionellen Repositorien wie bei PUB Publikationen an der Universität Bielefeld³ oder Deposit Once der Technischen Universität Berlin⁴.

² Beispielsweise von EconStor, dem Open-Access-Server der Deutschen Zentralbibliothek für Wirtschaftswissenschaften / Leibniz-Informationszentrum Wirtschaft (ZBW). EconStor dient der freien Veröffentlichung wissenschaftlicher Literatur aus den Wirtschaftswissenschaften. ZBW: EconStor <https://www.econstor.eu/dspace/> (Zugriff: 21.04.16)

³ Universität Bielefeld: Publikationen an der Universität Bielefeld <https://pub.uni-bielefeld.de/> (Zugriff: 21.04.16)

⁴ TU Berlin: Repository for Research Data and Publications <https://depositonce.tu-berlin.de/> (Zugriff: 21.04.16)

Ziele und Strategien für die Publikation von Forschungsdaten im Fachrepositorium Lebenswissenschaften

Die Publikationsmöglichkeiten für Forschungsdaten im Rahmen von PUBLISSO bauen auf dem strategischen Ziel auf, bereits vorhandene Infrastrukturen zur Datenpublikation in den Lebenswissenschaften aufzuzeigen und an den Stellen eigene Angebote aufzubauen, wo Lücken identifiziert werden. Dies bezieht sich beispielsweise auf den sog. long tail der Forschungsdaten, also Daten, die ein geringes Datenvolumen aufweisen, in verschiedenen Datenformaten vorliegen und somit nur schwer standardisierbar sind, aber auch auf lebenswissenschaftliche Teildisziplinen, in denen Möglichkeiten zur Datenarchivierung und -publikation weitgehend fehlen.

Als einen ersten Schritt zur Erfüllung dieses Ziels wird das Fachrepositorium Lebenswissenschaften zur Aufnahme von Forschungsdaten ausgebaut. Dabei werden wiederum folgende Ziele verfolgt:

- dauerhafte Archivierung von Forschungsdaten aus den Lebenswissenschaften,
- Sicherstellung der Nachnutzbarkeit der Forschungsdaten,
- Publikation singulärer Forschungsdaten sowie von Forschungsdaten, die mit einem Volltext verknüpft sind,
- Lizenzierung der Daten im Sinne von Open Data, also der Möglichkeit zur Nachnutzung.

Obwohl die Abdeckung des lebenswissenschaftlichen Fächerspektrums Medizin, Gesundheit, Ernährung, Umwelt und Agrar einen eher generischen Ansatz bedeutet, sollen auch fachspezifische Dateneigenschaften berücksichtigt werden.

Wesentliche Schritte bei dem Ausbau des Fachrepositoriums sind:

1. Entwicklung eines Metadatenschemas (Erfassungsmaske),
2. technische Implementierung der Erfassungsmaske,
3. Erprobung durch Publikation realer Forschungsdaten.

Der Beitrag beschäftigt sich mit dem ersten dieser Schritte.

Bei der Erarbeitung der relevanten Metadaten sollten darüber hinaus in Pflicht- und optionale Kriterien unterschieden werden.

Prüfung vorhandener Metadatenschemata

Als Orientierungshilfe für die Entwicklung eines Metadatenschema zur Erfassung von Forschungsdaten wurden zunächst Metadatenschemata existierender (Daten-) Repositorien analysiert und auf eine Übertragbarkeit für die vorliegende Aufgabenstellung geprüft. Folgende Quellen und Beispiele sind dabei einbezogen worden (Tab. 1).

Institution	relevante Inhalte	Fachgebiet
Digital Curation Center (DCC) ⁵	Auflistung von bereits erstellten Metadatenstandards diverser Disziplinen (<i>Grundausswahl</i>)	generisch
Humboldt Universität Berlin ⁶	Informationen zur Dokumentation von Forschungsdaten sowie zu geeigneten Metadatenschemata	generisch / Beispiele für Disziplinen
DataCite ⁷	Auswahl zwischen festgelegten und optionalen Inhalten (<i>Möglichkeit auf Diversität der Disziplinen einzugehen</i>)	generisch
GESIS Datorium ⁸	Beispiel für allgemeine Aufnahme von Forschungsdaten	Sozialwissenschaften
Pangaea ⁹	Datenrepositorium für Geowissenschaften	Geo-/ Agrarwissenschaften
Dublin Core Metadata Initiative (DCMI) ¹⁰	weltweiter Metadaten-Standard	
Darwin Core ¹¹	Weiterentwicklung von DCMI zur Darstellung von biologischen Spezifikationen	Biologie / Geologie
PUB Uni Bielefeld ¹²	Beispiel zur Darstellung von Forschungsdaten im Repository	generisch

Tabelle 1: Quellen und Beispiele für Metadatenschemata

⁵ Digital Curation Center: <http://www.dcc.ac.uk/resources/metadata-standards> (Zugriff 28.04.2016)

⁶ Humboldt Universität Berlin: <https://www.cms.hu-berlin.de/de/ueberblick/projekte/dataman/teilen/dokumentation> (Zugriff 28.04.2016)

⁷ Data Cite: http://schema.datacite.org/meta/kernel-3/doc/DataCite-MetadataKernel_v3.1.pdf (Zugriff 28.04.2016)

⁸ gesis: <https://datorium.gesis.org/xmlui/> (Zugriff 28.04.2016)

⁹ PANGAEA Datapublisher for Earth and Environmental Science: <https://www.pangaea.de/> (Zugriff 28.04.2016)

¹⁰ Dublin Core Metadata Initiative: <http://dublincore.org/documents/dcmi-terms/> (Zugriff 28.04.2016)

¹¹ Biodiversity Information Standards: <http://rs.tdwg.org/dwc/terms/index.htm#livingspecimenindex> (Zugriff 28.04.2016)

¹² PUB Universität Bielefeld: <http://pub.uni-bielefeld.de/data/2777409> (Zugriff 28.04.2016)

Eine der wichtigsten Quellen stellte das DataCite-Metadatenchema dar, das aktuell in der Version 3.1 vom Juni 2015 vorliegt¹³, im Laufe des Jahres 2016 jedoch aktualisiert wird. DataCite ist ein internationales Konsortium, das die Auffindbarkeit und Zitierbarkeit von frei zugänglichen Informationsressourcen, insbesondere Forschungsdaten, durch die Vergabe eines persistenten Identifiers fördert. ZB MED ist DataCite-Mitglied und fungiert als Vergabestelle für Digital Object Identifier (DOI) für lebenswissenschaftliche, akademische bzw. gemeinnützige Einrichtungen.

Auch die bisher bereits im Fachrepositorium Lebenswissenschaften veröffentlichten Informationen erhalten automatisch einen DOI.

Das in DataCite verwendete Metadatenchema hat keinen fachlichen Fokus, da die Referenzierung verschiedenster Objekte aus allen Disziplinen angestrebt wird. Es versucht jedoch auch fachspezifische Aspekte einzubinden, z.B. durch das Feld GeoLocation. Außerdem setzt es die Hürde zur Registrierung und damit Publikation von Forschungsdaten durch lediglich sechs verpflichtende Metadaten recht niedrig an und unterscheidet darüber hinaus in „empfohlene“ und „optionale“ Felder.

Pflichtfelder	Empfohlene Felder	Optionale Felder
Identifier (with type Attribut)	Subject (with scheme sub-property)	Language
Creator (with type and name identifier attributes)	Contributor (with type, name identifier, and affiliation sub-properties)	Size
Title (with optional type attribut)	Date (with type sub-property)	Format
Publisher	RelatedIdentifier (with type and relation type sub-properties)	Version
Publication Year	Description (with type sub-property)	Rights
ResourceType (with general type description sub-property)	GeoLocation (with point and box sub-properties)	

Tabelle 2: aktuelles Metadatenchema DataCite

¹³ Data Cite: http://schema.datacite.org/meta/kernel-3/doc/DataCite-MetadadataKernel_v3.1.pdf (Zugriff 28.04.2016)

**Metadatenchema für Forschungsdaten im Fachrepositorium
Lebenswissenschaften**

Insbesondere auf dem DataCite-Schema, aber auch auf den anderen Beispielen sowie den Anforderungen der Forschungsdaten in den lebenswissenschaftlichen Disziplinen fußt das für das Fachrepositorium Lebenswissenschaften entwickelte Metadaten-chema:

Metadatum (übergeordneter Begriff)	Feldname	Feldname (untergeordnet)
Titel	Titel	
Urheberschaft	Autor*in	Nachname
		Vorname
		Orcid (optional)
		Affiliation (optional)
	Körperschaft (wenn kein Autor vorhanden)	
Dateiupload	Hochzuladende Datei	
	Format (xls, jpeg, etc.)	
	Medientyp (Bild, Video, Software, etc.)	
	Größe	
	Zugriffsrechte (open access, Embargo)	Embargofristende
	Copyrightjahr	
	Lizenz	
	DOI	Neu
		Vorhanden
Zuletzt hochgeladen		
Erschließung	Abstract	Sprache
	Fachgruppenzuordnung	
	DDC-Klassifikation	
	Sprache	

Tabelle 3: Metadatenchema für Forschungsdaten im Fachrepositorium Lebenswissenschaften (Pflichtfelder)

Metadatum (übergeordneter Begriff)	Feldname	Feldname (untergeordnet)
Beteiligte	Beteiligte Personen	OrclD
		Affiliation
	Förderer	Förder-ID
Erfassung	Schlagworte	Sprache
	Datenerhebungsform	
	Erhebungszeit	Zeitpunkt
		Zeitraum
	Erfassungsort	Koordinaten (Point)
Kordinaten (Box)		
Externe Referenzen	Verwendete Publikationen	
	Zugehörige Publikationen	
	Versionen	Vorgänger
		Nachfolger

Tabelle 4: Metadatenschema für Forschungsdaten im Fachrepositorium
Lebenswissenschaften (optionale Felder)

Dieses Schema war das Ergebnis eines längeren Entwicklungs- und Diskussionsprozess bei ZB MED. Die wichtigsten Punkte dieses Prozesses werden im Folgenden dargestellt.

Format und Medientyp

Auch wenn nicht alle Formate entsprechend im Fachrepositorium wiedergegeben werden können, ist es für die Nutzerinnen und Nutzer wichtig zu wissen welches Dateiformat vorliegt, um beurteilen zu können, ob eine entsprechende Wiedergabemöglichkeit vorliegt. Beispiele für Dateiformate von Forschungsdaten sind PDF, JPEG, XLS oder MPEG. Wichtig wird diese Angabe vor allem auch dann, wenn ZIP-Dateien, also Datencontainer, abgelegt werden.

Der Medientyp geht in der Regel mit einem bestimmten Dateiformat einher, ist jedoch als weiteres Feld aufgenommen worden, da Forschungsdaten sehr unterschiedlich definiert sein können und die Nutzerinnen und Nutzer dann sofort unterschiedliche Typen erkennen können. Hinzu kommt, dass beispielsweise für Audio- und Videodateien HTML-basierte Player direkt im Fachrepositorium zur Verfügung stehen. Als Medientypen können Audio, Video, Software, Text, Bild oder weitere angegeben werden.

Zugriffsrechte / Möglichkeit eines Embargos für die Publikation

Die Definition von Zugriffsrechten bietet die Auswahl zwischen „Open Access“ und „Embargo“. Eine Publikation der eigentlichen Daten ist demzufolge auf jeden Fall vorgesehen. Der Autor kann jedoch eine Embargofrist setzen und somit für eine gewisse Zeit lediglich die Metadaten für seine Publikation sichtbar machen.

Die Möglichkeit für das Setzen einer Embargofrist gründet auf einem wichtigen Vorbehalt der Forschenden in Bezug auf eine Datenpublikation: die Befürchtung, dass andere Wissenschaftler die Daten auswerten und Ergebnisse veröffentlichen bevor der Autor seine eigenen Forschungen abgeschlossen hat und ihm ggf. Möglichkeiten der Erstpublikation vorweggenommen werden (Fecher, B. Friesike, S., Hebing, M., Linek, S., Sauermann, A., 2015).

Vergabe von Lizenzen

Da die Datenpublikation im Fachrepositorium Lebenswissenschaften im Sinne von Open Data die Möglichkeit einer Nachnutzung impliziert, muss eine entsprechende Lizenz dafür vom Autor vergeben werden.

Obwohl Forschungsdaten vom Grundsatz her nicht urheberrechtlich geschützt sind, stellen sie im juristischen Sinne und im wissenschaftlichen Kontext vielfach ein sog. Datenbankwerk dar, da vom Forschenden eine „wesentliche Bearbeitung“ im Sinne einer systematischen Aufbereitung erfolgt.

Für Forschungsdaten kommt in der Regel die Open Data Commons Open Database License (ODbL)¹⁴, für Software die General Public Licence GNU¹⁵ in Frage.

Weit verbreitet für Textpublikationen, Präsentationen, Poster, etc. sind die Creative-Commons-Lizenzen¹⁶.

Da insbesondere hinsichtlich der Lizenzvergabe bei Forschungsdaten bei Autorinnen und Autoren eine große Unsicherheit besteht, stellt sich für die Publikation im Fachrepositorium Lebenswissenschaften die Frage, ob eine Empfehlung ausgesprochen werden soll oder die Forschenden selbst eine Auswahl treffen. Neben einer Auswahlmöglichkeit in Frage kommender Lizenzen für den publizierten Medientyp, erleichtert eine Empfehlung sicherlich die Publikation für den Autor.

¹⁴ Open Data Commons: <http://opendatacommons.org/licenses/odbl/> (Zugriff 26.04.2016)

¹⁵ Free Software Foundation: <http://www.gnu.org/licenses/gpl.html> (Zugriff 26.04.2016)

¹⁶ Creative Commons Deutschland: <http://de.creativecommons.org/was-ist-cc/> (Zugriff 26.04.2016)

Bevorzugung des DOI als persistenten Identifikator

Da ZB MED als DataCite-Mitglied DOIs als persistente Identifikatoren vergibt, verwendet es ihn auch in den eigenen Services. Ausgehend davon, dass es sich bei den künftigen Veröffentlichungen von Forschungsdaten um Erstpublikationen handelt, wird daher automatisch ein DOI vergeben. Sollte für Forschungsdaten bereits ein persistenter Identifikator existieren, wird dieser übernommen, gleichgültig ob es sich um einen DOI, eine URN oder ein Handle handelt. Dies muss allerdings manuell erfolgen.

Einbindung der DDC für die Sacherschließung

Neben der Fachgruppenzuordnung, die die lebenswissenschaftlichen Fächer Medizin, Gesundheit, Ernährung, Agrar, Umwelt und Biologie zur Auswahl stellt, war das Ziel einer Einbindung von Thesauri oder Klassifikationen zur Ableitung von Schlagworten für das Retrieval. Naheliegender ist die Einbindung der Dewey Decimal Classification (DDC) als weit verbreiteten Standard für eine detailliertere fachliche Zuordnung. Insbesondere bei digitalen Publikationen greift die Möglichkeit einer Mehrfachangabe von fachlichen Zuordnungen, um den Sucheinstieg auszuweiten. In der Praxis stellt sich dies für alle lebenswissenschaftlichen Fächer jedoch als schwer durchführbar heraus, da hier bis zu 35 Klassifikationen relevant sind. Eine richtige Zuordnung fällt daher schwer und aus Gründen der Qualitätssicherung wurde die Auswahl auf die Hauptsachgruppen 610 (Medizin, Gesundheitswesen), 610 & 640 (Ernährungswissenschaft), 630 (Agrarwissenschaft) und 333.7 (Umwelt) beschränkt.

Zulassung verschiedener Sprachen

Zu der Frage in welchen Sprachen die Publikation von Forschungsdaten erlaubt wird, wird eine recht weite Lösung angestrebt und im europäischen Kontext die fünf Sprachen Englisch, Deutsch, Spanisch, Französisch und Italienisch zugelassen. Das Argument dafür ist in erster Linie, dass die Forderung einer Übersetzung ins Englische oder Deutsche eine hohe Barriere für die Forschenden aufbaut. Das Datenrepositorium Dryad¹⁷ beispielsweise erlaubt lediglich die Ablage englischsprachiger Forschungsdaten.

¹⁷ Dryad Digital Repository: <http://datadryad.org/pages/policies> (Zugriff 28.04.2016)

Die Erfahrung in der Kooperation mit Dryad im Rahmen des Publikationsportals German Medical Science (GMS) von ZB MED, das die Möglichkeit einer Datenpublikation ergänzend zu einem Zeitschriftenartikel kostenfrei eröffnet, bestätigt dies. Die Zulassung mehrerer Sprachen bedeutet aber auch, dass Abstracts in allen fünf Sprachen akzeptiert werden. Ein englischer Abstract soll jedoch auf jeden Fall vorliegen. Dies ist bei Textpublikationen durchaus eine übliche Vorgehensweise. Inwieweit eine Einbindung von Übersetzungs-Tools möglich wäre, ist eine Frage, die künftig geklärt werden sollte, zum Start des Repositoriums jedoch keine Priorität hat.

Datenerhebungsform

Die Datenerhebungsformen, also die Art und Weise wie Daten gewonnen wurden, sollen anhand einer Dropdown-Liste wählbar sein. Das Problem hierbei ist, dass es kaum Vergleichsmöglichkeiten aus anderen Repositorien gibt und der Umfang aufgrund der Vielfalt in den Lebenswissenschaften schwer zu ermitteln ist. Eine Auswahl möglicher Erhebungsformen stellt die aktuelle Liste dar: Interview, Umfrage, Anamnese, Exploration, Probe, Gewebeprobe, Flächenmischprobe, Bodenbohrung, Experiment, apparative Untersuchung, körperliche Untersuchung, Feldbeobachtung, Laborbeobachtung, Analyse, Genomsequenzierung, Messung, Berechnung, Evaluation, Querschnittstudie, Langzeitstudie, Interventionsstudie, Kohortenstudie, Simulation, andere.

Verwendete Publikationen

Bei einer Erstpublikation von Forschungsdaten ist es derzeit eher unwahrscheinlich, dass sie auf anderen publizierten Daten aufbauen bzw. diese benutzen. Das Feld kann aber in Zukunft relevant werden, wenn das Aufkommen von publizierten Forschungsdaten steigt.

Ausblick

Das dargestellte Metadatenschema muss in einem nächsten Schritt technisch in das System integriert werden. Hierbei und bei der praktischen Erprobung im Rahmen der Publikation von Forschungsdaten wird sich zeigen, inwieweit die als „ideal“ oder „wünschenswert“ eingestuften Kriterien beibehalten werden können, welche wegfallen oder welche ergänzt werden müssen.

Neben dem Metadatenschema werden bei der Umsetzung noch weitere Faktoren eine Rolle spielen. Beispielsweise ist zu berücksichtigen, dass bei der Einreichung der Daten nicht notwendigerweise alle erforderlichen Informationen vorliegen, so dass es einer Rücksprache mit den Autorinnen oder Autoren bedarf. Hierfür sind entsprechende Prozesse zu entwickeln und Personalkapazität vorzuhalten.

Dies und andere Punkte werden sich erst in den weiteren Phasen des Prozess ergeben.

Literaturhinweise

Quast, Andres: ZB MED Fachrepositorium Lebenswissenschaften - Open Access im Spiegel veränderter digitaler Publikationsformen. - In: Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen: Fakten und Perspektiven 2015. Jahresbericht. - 2016, in Press.

Fecher, B. Friesike, S., Hebing, M., Linek, S., Sauermann, A.: A Reputation Economy: Results from an Empirical Survey on Academic Data Sharing. Deutsches Institut für Wirtschaftsforschung, Discussion Paper 1454, 2015, S. 4-5