

Erprobung eines Metadatenmodells zur Beschreibung von FDM-Services

Lemaire, Marina

marina.lemaire@uni-trier.de
Universität Trier, Deutschland
ORCID: 0000-0003-4726-2481

Christ, Andreas

christ@ub.uni-kiel.de
Christian-Albrechts-Universität zu Kiel, Deutschland
ORCID: 0000-0002-3591-2355

Das Management von Forschungsdaten (FDM) hat für Wissenschaftler*innen zunehmend an Bedeutung gewonnen. Deshalb suchen sie nach FDM-Services, die sie dabei unterstützen, effizient und qualitätsgesichert mit ihren Forschungsdaten umzugehen. Das Angebot und die Entwicklung solcher Dienste erfolgt auf vielen Ebenen. Es reicht von lokalen (institutionellen) Stellen, die an Hochschulen und außeruniversitären Forschungseinrichtungen zu finden sind, über landes- und bundesweit übergreifende Einrichtungen, wie z. B. FDM-Landesinitiativen oder die NFDI-Konsortien bis hin zu internationalen Angeboten, bspw. die European Open Science Association (EOSC) oder das Social Sciences and Humanities Open Cloud Projekt (SSHOC). Das Angebot ist vielfältig, auf verschiedenen Plattformen und/oder einfachen Webseiten zu finden. Zudem werden die Informationen sehr divers hinsichtlich Struktur und Informationsgehalt zur Verfügung gestellt. Für Forschende ist es daher einerseits oftmals problematisch, die auf ihre Bedarfe passenden Dienste zu finden und andererseits auf der Basis der bereitgestellten Informationen zu entscheiden, ob oder welcher Dienst zu ihren Bedarfen passt. Gleichzeitig stellt sich das Problem auf der anbietenden Seite, welche Informationen wie bereitgestellt werden müssen und wie man dafür sorgen kann, dass die Forschenden das eigene Angebot einerseits finden, aber auch verstehen bzw. erfassen können, ob es für sie passend ist und worin es sich im Vergleich zu einem ähnlichen Dienst unterscheidet. Aus der Komplexität der FDM-Bedarfe sowie der individuellen und heterogenen Organisation von FDM-Services an unterschiedlichen Einrichtungen ergeben sich grundlegende konzeptionelle Fragen für Serviceeinrichtungen bei der Bereitstellung von Dienstleistungsportfolios im Bereich des Forschungsdatenmanagements. Um diese Bedarfe zu adressieren, erscheint eine Plattform sinnvoll, auf der FDM-Services nach einem einheitlichen Schema beschrieben sind und Suchoptionen angeboten werden, die das schnelle und zielgerichtete

Auffinden der passenden Services ermöglicht. Ehe jedoch entsprechende technische Lösungen implementiert werden, sollte sich über ein Metadatenschema verständigt werden, damit einerseits die Services leichter vergleichbar werden und ein allgemeines Grundverständnis/Vokabular zur Bezeichnung und Beschreibung von FDM-Diensten sich leichter etablieren kann. Andererseits ist es sinnvoll, FDM-Servicebeschreibungen aus dem eigenen System in andere Plattformen leicht überführen zu können, um die Auffindbarkeit über Metasuchmaschinen zu erhöhen. Es gibt einige Plattformen, die Suchen zu FDM-Services allgemein oder zu spezifischen FDM-Servicetypen ermöglichen, wie z.B. re3data „Registry of Research Data Repositories“ (IBI et al. 2024), der ELN-Finder (ULB TU DA/ZB MED 2024), der Dienstekatalog der DHD AG Geisteswissenschaftliches Forschungsdatenmanagement (Rau / Helling 2022) oder der RDM Navigator (KIT 2021). Viele dieser Beschreibungsmodelle decken jedoch oft nur bestimmte Teilbereiche der gesamten FDM-Servicelandschaft ab, haben einen spezifischen institutionellen Blick und/oder sind auf bestimmte Fachrichtungen oder Servicetypen beschränkt. Gleichzeitig fehlt eine einheitliche, formale Beschreibungssprache für viele Kategorien in diesen Modellen, sodass Gleiches mit dem gleichen Vokabular benannt wird. Zusätzlich existiert noch kein klares Verständnis darüber, was genau ein FDM-Service ist. Die NFDI-Konsortialversammlung definiert einen Dienst als „eine technisch-organisatorische Lösung [...], die Speicher- und Rechenleistungen, Software, Prozesse und Workflows ebenso umfassen kann wie die notwendige personelle Betreuung für unterschiedliche Service-Desks“ (NFDI 2022). Bei IT-Service-Dienstleistern ist ITIL (Information Technology Infrastructure Library) ein verbreitetes Framework für das Management von IT-Services. Es definiert Service sehr viel generischer als ein „Bündel von spezialisierten, organisatorischen Fähigkeiten“, die einen Mehrwert bei Kund*innen erbringen. Die Beschreibung eines Dienstes in einem Servicekatalog ist ein elementarer Bestandteil des Service-Managements, aber deren genaue Ausgestaltung, insbesondere für die Kund*innen, muss individuell an die Zielgruppe angepasst werden. (Ebel 2021) D. h., dass ITIL zwar ein Grundgerüst für ein Service-Metadatenmodell bereitstellt, aber die zentralen FDM-spezifischen Fragen nicht beantwortet. Des Weiteren stellt sich die Frage, wie sich ein Dienst von einem anderen abgrenzt: Ist z. B. die Schulung zu einer Software ein eigenständiger Dienst neben der Bereitstellung der Software selbst, oder ist sie ein integraler Bestandteil?

Diesen Fragen widmet sich derzeit eine kleine Untergruppe aus der GO UNITE! Arbeitsgruppe FDM-Beschreibungsmodelle (<https://go-unite.de/index.php/ag-fdm-beschreibungsmodell/>) und hat einen ersten Vorschlag für ein Metadatenmodell zur Beschreibung von FDM-Services erarbeitet.¹ Es wurden verschiedene FDM-Servicebeschreibungen² aufeinander gemappt, die Gesamtübersicht aller Kategorien wurden in der Gruppe und in einem kleineren Workshop mit weiteren Interessierten diskutiert. Dabei wurden die Kategorienbezeichnungen sowie -beschreibungen

gen, der Datentyp (z. B. Text, Liste), die Kardinalität (1-1, 1-n, 0-1, 0-n), die öffentliche Sichtbarkeit und in einem Feld Beispiele und Erläuterungen bzw. Fragen notiert, bei denen in der Diskussion, z. B. noch nicht zu einer Einigung gefunden wurde. Insbesondere hinsichtlich des anzuwendenden Vokabulars in einigen Kategorien bleiben aufgrund der Komplexität des FDM einige Fragen offen. Die Autor*innen gehen bewusst den Schritt, diesen ersten Vorschlag mit den identifizierten Diskussionspunkten zu veröffentlichen, um mit einem größeren Kreis diese Aspekte zu erörtern und anhand der Ergebnisse, das Modell schrittweise gemeinsam mit der Community weiterzuentwickeln.

Der hier vorgeschlagene Workshop der DHd AG Geisteswissenschaftliches Forschungsdatenmanagement (gwFDM) setzt genau dort an und möchte die Diskussion des Modells beginnen. Denn die Erprobung eines Modells gelingt insbesondere dann, wenn versucht wird, das Schema auf existierende FDM-Services anzuwenden, um einerseits sich den bereits identifizierten Fragen zu widmen und hoffentlich Antworten darauf zu finden; andererseits um weitere Probleme in der Anwendung des Schemas zu identifizieren und dafür Lösungen zu erarbeiten. Viele Mitglieder der AG gwFDM stellen in ihrer Funktion als Dienstleister (Datenzentren, FDM-Servicestellen und NFDI-Konsortien) FDM-Services bereit und haben bereits vor einigen Jahren einen Dienstekatalog (Rau/Helling 2022) erstellt. Dieser bietet aufgrund seiner Heterogenität, eine hervorragende Grundlage: In ihm sind Dienste aus unterschiedlichen Datenzentren und anderen FDM-Service-anbietenden Einrichtungen enthalten, die ein breites Spektrum an Servicetypen und Beschreibungen aufweisen, sodass eine sehr große Spannbreite an FDM-Services für die Evaluation des Modells zur Verfügung stehen. Da bei der Entwicklung des Modells eben nicht nur die Anbieterseite, sondern auch die Kund*innenperspektive berücksichtigt werden soll, bietet das Publikum der DHd ein gutes Spektrum an Expertisen und Sichtweisen für die Diskussion.

Um das Metadatenmodell im Rahmen des Workshops zu evaluieren, vorhandene Herausforderungen zu diskutieren und gemeinsam Lösungsstrategien zu entwickeln, ist folgender Ablauf geplant: Nach der Begrüßung und einer Vorstellungsrunde der Teilnehmenden beginnen wir mit der einführenden Vorstellung des Metadatenmodells sowie der Darlegung der offenen Fragen, die bereits in der Publikation benannt wurden. Anschließend werden die Teilnehmenden in kleine Gruppen aufgeteilt, um das Modell praktisch auf von ihnen ausgewählte Services des Dienstekatalogs anzuwenden. Ziel ist es, Erfahrungen zu sammeln, potenzielle Verbesserungsvorschläge zu erarbeiten und ggf. weitere Herausforderungen zu identifizieren. Nach einer kurzen Pause präsentieren die Gruppen ihre Ergebnisse und alle Workshopeteilnehmenden führen eine moderierte Diskussion über die Herausforderungen, die während der Anwendung des Metadatenmodells auftraten sowie über ihre ggf. bereits angewandten Lösungsstrategien. In dieser Diskussion sollen weitere Lösungsansätze entworfen werden. Alle Aspekte der Diskussion werden auf einer Mindmap gesammelt, um sie anschließend als Grundlage in der nächsten

Workshop-Phase zur Verfügung zu haben. Daran anschließend werden in der Konsolidierungsphase die Erkenntnisse durch das Clustern der Mindmap zusammengefasst und die nächsten Schritte für die Weiterentwicklung des Metadatenmodells diskutiert und festgelegt.

Ablaufplan für einen halbtägigen Workshop

| | |
|---|--------------|
| Begrüßung und Einführung | (20 Minuten) |
| Präsentation des FDM-Service-Metadatenmodells | (25 Minuten) |
| Gruppenarbeit: Erprobung des Metadatenmodells anhand des Dienstekatalogs | (1 Stunde) |
| Pause | (15 Minuten) |
| Präsentation und Diskussion der Gruppenergebnisse | (1 Stunde) |
| Pause | (15 Minuten) |
| Konsolidierung der Diskussionsergebnisse und Festlegung der nächsten Schritte | (45 Minuten) |

Auf diese Weise bietet der Workshop eine interaktive Plattform für alle Beteiligten, die an der Optimierung von Metadatenmodellierung für FDM-Services interessiert sind. Das Metadatenmodell wird durch die Bewertung seiner Effektivität und Vollständigkeit bei der Beschreibung verschiedener FDM-Services evaluiert. Dabei wird gezielt nach Lücken und Verbesserungspotenzialen durch das Sammeln von konstruktivem Feedback zur Verfeinerung des Modells gesucht. Zusätzlich fördert der Workshop ein erweitertes Verständnis bei den Teilnehmenden für die Vielfalt und Komplexität von FDM-Services. Die unterschiedlichen Erfahrungshorizonte der Teilnehmenden tragen zu einer reichhaltigen und sich gegenseitig befruchtenden Diskussion bei, die es ermöglicht, umsetzbare Erkenntnisse zu generieren und konkrete Empfehlungen für die Weiterentwicklung und Verbesserung des Metadatenmodells zu erhalten, um dessen Implementierung in der Breite der FDM-Servicelandschaft zu begünstigen. Die Ergebnisse des Workshops werden in einem Blogbeitrag publiziert. Des Weiteren könnte dieser Workshop ein Startschuss sein für die Überarbeitung und Aktualisierung des Dienstekatalogs der AG gwFDM. Darüber hinaus ist er der Auftakt für die weitere Zusammenarbeit mit der Go Unite! AG in diesem Feld und darüber hinaus.

Fußnoten

1. Die Autor*innen sind Patrick Helling, Boris Jakob, Stefan Kellendonk, Marina Lemaire und Rouven Schabiner. (Patrick Helling und Marina Lemaire sind Mitglieder der DHd AG Geisteswissenschaftliches Forschungsdatenmanagement.) Der Entwurf für das Metadatenmodell ist bereits fertiggestellt. Derzeit werden die Rahmentexte überarbeitet. Die Publikation wird bis Ende dieses Jahres erscheinen. Alle Autor*innen sind mit der Einreichung dieses Abstracts vor der Publikation einverstanden.
2. Die verwendeten Schemata sind: Dienstekatalog der DHd AG Geisteswissenschaftliches Forschungsdatenmanagement (Rau / Helling 2022), DIAMANT-Ser-

vice-Beschreibung (Lemaire et al. 2020), RDMNavigator (KIT 2021), Serviceverzeichnis Forschungsdaten (Erven et al. 2019), Referenzmodell für Strategieprozesse im institutionellen Forschungsdatenmanagement (RISE-DE) (Hartmann/Jacob/Weiß 2019) und EOSC EU Node - Resource Hub (Sanchez-P. 2021).

Bibliographie

Ebel, Nadin (2021): *Basiswissen ITIL 4: Grundlagen und Know-how für das IT Service Management und die ITIL-4-Foundation-Prüfung*. 1. Aufl. Dpunkt.Verlag GmbH.

Erven, Judith / Dierkes, Jens / Aguilera, Alvaro / Brand, Ortrun / Ludwig, Jens / Müller-Pfefferkorn, Ralph / Schubert, Paul / Sutter, Paul (2019): *Serviceverzeichnis Forschungsdaten*. <https://doi.org/10.11588/heibooks.598.c8436>.

Hartmann, Niklas K. / Jacob, Boris / Weiß, Nadin (2019): *RISE-DE – Referenzmodell für Strategieprozesse im institutionellen Forschungsdatenmanagement. Version 1.0*. Zenodo. <https://doi.org/10.5281/zenodo.3585556>.

IBI, Institut für Bibliotheks- und Informationswissenschaft / KIT Bibliothek, Bibliothek des Karlsruher Instituts für Technologie / OS Helmholtz, Helmholtz Open Science / Lib Purdue, Libraries of Purdue University (2024): *re3data.org. – Registry of Research Data Repositories*. <http://www.re3data.org/> (letzter Zugriff 22. Juli 2024).

KIT, Karlsruher Institut für Technologie (2021): *RDM Navigator*. <https://www.rdm.kit.edu/RDM/RDM-navigator/index.php> (letzter Zugriff 22. Juli 2024).

Lemaire, Marina / Gerhards, Lea / Kellendonk, Stefan / Blask, Katarina / Förster, André (2020): *Das DIAMANT-Modell 2.0. Modellierung des FDM-Referenzprozesses und Empfehlungen für die Implementierung einer institutionellen FDM-Servicelandschaft*. (= eSciences Working Papers 05) Trier.

NFDI, Konsortialversammlung des Vereins Nationale Forschungsdateninfrastruktur (NFDI) e.V. (2022): *Stellungnahme der NFDI-Konsortien zu Basisdiensten*. Zenodo. <https://doi.org/10.5281/zenodo.6091657>.

Rau, Felix / Helling, Patrick (2022): "Der Dienstekatalog der AG Datenzentren - Ein digitales Verzeichnis für Forschungsdatenmanagement-Services in den Geisteswissenschaften", 10.5281/ZENODO.6328019.

Sanchez-P., Jorge-A. (2021): "EOSC Portal Profiles v4.00", 10.5281/zenodo.5726890.

ULB TU DA, Universitäts- und Landesbibliothek Darmstadt / ZB MED, Informationszentrum Lebenswissenschaften (2024): *ELN-Finder*. <https://eln-finder.ulb.tu-darmstadt.de/home> (letzter Zugriff 22. Juli 2024).