

Metadaten-Werkzeuge für die NFDI (Metadata-Tools4NFDI)

Sprecher: Prof. Dr. Thomas Stäcker, ULB Darmstadt, thomas.staecker@tu-darmstadt.de

Welches Forschungsgebiet soll adressiert werden?

Metadata-Tools4NFDI richtet sich als Basiskonsortiums-Initiative an alle Fächer.

Mit welcher Art von Daten beschäftigen Sie sich?

Metadata-Tools4NFDI will Werkzeuge für die leichte und bestmöglich automatisierte Erzeugung von detailreichen und interoperablen Metadaten durch Forschende bereitstellen. In der Anwendbarkeit auf spezifische Datentypen oder -domänen gibt es keine Einschränkungen.

Welche wesentlichen Maßnahmen planen Sie, um Datenmanagement für Ihr Forschungsgebiet einzuführen, und welche Services wollen Sie anbieten?

Unser angestrebtes Arbeitsprogramm ist wie folgt skizziert, es besteht aus dem Zusammenfügen und Bereitstellen von Vorhandenem und der ergänzenden Entwicklung von Werkzeugen für identifizierte Lücken. In der Diskussion werden sich ggf. noch weitere Aspekte oder in Einzelcommunities vorhandene Werkzeuge ergeben. Es ist unser Ziel diese Werkzeuge und ganze Werkzeugketten fachunabhängig verfügbar und nutzbar zu machen.

1) Werkzeuge für die Standardisierung von Metadaten

1a) Generatoren / Frameworks / Workflows zur Terminologieerstellung

1b) Generatoren und Plattformen für die gemeinsame Nutzung und Wiederverwendung von Metadatenschemata

1c) Kataloge oder Verzeichnisse für Metadatenschemata und Terminologien

2) Werkzeuge für den Umgang mit Metadaten in der aktiven Forschung - Daten-Organisation durch semantisches tagging

Idee: Angereicherte Verzeichnisstruktur als "universelles" ELN. Es gibt nicht das eine ELN für alle, viele Anwender können in absehbarer Zeit kein für ihre Bedürfnisse passendes ELN erwarten. Verzeichnisstrukturen hingegen sind gängige Praxis für die Datenverwaltung, wenngleich es ihnen an Standardisierung und Komfort mangelt. Letzteres könnte durch die Anreicherung einer Verzeichnisstruktur mit strukturierten Metadaten überwunden werden.

Ziel: Angebot einer Umgebung zum Anhängen und Verwalten von kontrollierten Tags (semantischen Tripeln) an die Verzeichnisstruktur. Könnte ein Kommandozeilen- oder GUI-Ansatz sein. Die Tags werden verwendet für

- Die Indexierung zum Auffinden von Verzeichnissen
- Komfort-Routinen für Aufgaben wie
 - Sicherung und Wiederherstellung
 - Übertragung von und zu einem entfernten Ort
 - Veröffentlichung im Repository
- Bereitstellung von maschinell verwertbaren Metadaten für Aufgaben wie Analyse oder Visualisierung

3) Werkzeuge für die Erzeugung von Metadaten

3a) Harvesting von Dateien mit Metadaten-Informationen zur Erstellung von kontrollierten Metadaten (z.B. Tags für die Verzeichnisstruktur oder Metadatenschema für die Publikation)

3b) Synergetische Nutzung der identischen Metadatenstruktur für andere Anwendungen

Angebot von Software-Bibliotheken für die Interaktion mit kontrollierten Metadaten aus typischen Anwendungen, die in der Forschung eingesetzt werden. Optionen für die synergetische Nutzung sind

- Prozesssteuerung
- Laborautomatisierung
- Analyse und Visualisierung

3c) ELN-Integration

4) Werkzeuge für die Veröffentlichung

4a) Plattformen für die Veröffentlichung von fachspezifischen Metadatensätzen

4b) Anbindung an Initiativen und Konzepte wie FAIR digital objects

5) Werkzeuge für Data Science

5a) Mappings zwischen Standardisierungen (Terminologien oder Schemata)

5b) Einstiegspunkte für Abfragen

Über den W3C-Standard „Linked Data Plattform“ (LDP) lassen sich die Metadatensätze (und Knowledgegraphen allgemein) austauschen.

6) Nutzbarmachung der Werkzeuge durch Trainings-, Fortbildungs- & Outreach-Aktivitäten

Welche speziellen Anforderungen an das Forschungsdatenmanagement gibt es in Ihrem Gebiet und wie wollen Sie diese adressieren?

n/a

Welche Erfahrungen/welchen Hintergrund hat Ihre Gruppe in Datenmanagement?

Die Akteure unserer Initiative sind innerhalb und außerhalb der NFDI in mehreren einschlägigen Projekten in der Erzeugung von Metadaten-Werkzeugen aktiv, z.B.:

- DFG-Projekt AIMS („Applying Interoperable Metadata Standards“) (ITC Aachen, ULB Darmstadt)
- FAIR Data Space („FAIRness Monitoring“ und „Terminologies & Validation“) (ITC Aachen)
- Internes Projekt „Coscine“ und weitere Kooperationsprojekte mit i5 / Fraunhofer FIT (ITC Aachen)
- Projekte „Dipl-Ing“ (BMBF) und „ReSUS“ (DFG) (UB Stuttgart)
- DFG-Projekt MASi „Metadata Management for Applied Sciences“ (KIT, TU Darmstadt)
- Helmholtz Metadata Collaboration (HMC) Platform (KIT)
- MetaStore “Metadata Repository and Schema Registry” (KIT)
- NFDI4Ing, v.a. Measure S-3 (alle Antragssteller)

Welche relevanten (internationalen) Partner und vorhandenen Infrastrukturen möchten Sie zusammenbringen?

Wir möchten alle bestehenden und zukünftigen NFDI-Konsortien und die dahinter stehenden Infrastrukturpartner zusammenbringen um gemeinsam an einem interoperablen Set an

Werkzeugen zur Metadaten-Generierung zu arbeiten, das natürlich auch über Deutschland / die NFDI hinaus genutzt werden kann und soll. Internationale Bezugspunkte sind dabei neben der EOSC auch die entsprechenden RDA-WGs (z.B. I-ADOPT, Metadata Standards Catalog, Research Metadata Schemas).

Wo sehen Sie Schnittstellen zu der gesamten NFDI?

Das Thema wie Forschende überhaupt qualitätvolle, detailreiche und interoperable Metadaten erzeugen können, ist immanent wichtig für den Erfolg großer Teile der NFDI. Ohne leicht bedienbare Werkzeuge dazu wird eine Verankerung in der Breite der Forschung nur schwer gelingen.

Besondere Nähe sehen wir bei diesen bestehenden Konsortien:

- NFDI4Ing, NFDI4BioDiversity, NFDI4Chem, NFDI4Culture, DataPlant
- sowie aus der zweiten Runde:
- FAIRmat, NFDI4Earth, NFDI-MatWerk, PUNCH

Welche Querschnittsthemen erachten Sie für Ihr Konsortium und für die NFDI insgesamt als relevant?

Zu welchen dieser Querschnittsthemen könnte Ihr Konsortium wie beitragen?

Unsere Konsortiumsinitiative adressiert von den bislang identifizierten Querschnittsthemen v.a. „Metadata, Findability“, „Terminologies“ und „Provenance“, aber auch „Infrastructure, Interoperability, Interfaces“. Mithin also die NFDI-Sektionen „Metadata, Terminologies, Provenance“ und auch „Common Infrastructures“.

Was erwarten Sie von Ihrer Teilnahme an der NFDI-Konferenz?

Wir möchten das aus unserer Sicht sehr wichtige Thema der Metdaten-Tools in die Konferenz tragen und unsere Ideen dazu zu einem Baustein für ein umfassendes Basisdienste-Konsortium machen, an dessen Entwicklung wir uns gerne beteiligen wollen.

Vorgesehene Mitglieder des Konsortiums (Co-Sprecherinnen/Co-Sprecher und die weiteren, beteiligten Institutionen):

Co-Sprecher/in	Zugehörige Institution
Prof. Dr. Thomas Stäcker Direktor der Universitäts- und Landesbibliothek thomas.staecker@tu-darmstadt.de	Technische Universität Darmstadt Magdalenenstr. 8 64289 Darmstadt
Prof. Dr. Matthias Müller Leiter des IT-Centers mueller@itc.rwth-aachen.de	RWTH Aachen University Seffenter Weg 23 52074 Aachen
Dr. Helge Steenweg Direktor der Universitätsbibliothek helge.steenweg@ub.uni-stuttgart.de	Universität Stuttgart Holzgartenstr. 16 70174 Stuttgart
Dr. Irina Sens Leitung Bibliotheksbetrieb irina.sens@tib.eu	Technische Informationsbibliothek – Leibniz-Informationszentrum Technik und Naturwissenschaften Welfengarten 1 B 30167 Hannover
Prof. Dr. Achim Streit Direktor des Steinbuch Centers for Computing achim.streit@kit.edu	Karlsruher Institut für Technologie Hermann-von-Helmholtz-Platz 1 76344 Eggenstein-Leopoldshafen
Prof. Dr. Britta Nestler Professorin für Mikrostruktursimulation in der Werkstofftechnik britta.nestler@kit.edu	Karlsruher Institut für Technologie Institut für Angewandte Materialien / Computational Materials Science Straße am Forum 7 76131 Karlsruhe