Section

(Meta)data, Terminologies, Provenance

Working Group Charter
Ontology harmonization and mapping

Name of the working group

Ontology harmonization and mapping

Acronym

section-metadata-wg-onto

Contact persons

Holger Israel (Technische Informationsbibliothek)
holger.israel@tib.eu

Lozana Rossenova (Technische Informationsbibliothek; NFDI4Culture)
lozana.rossenova@tib.eu

Roman Baum (Deutsche Zentralbibliothek für Medizin; NFDI4Health)
baum@zbmed.de

Authors

Ivonne Anders, Tobias Arera-Rütenik, Susanne Arndt, Roman Baum (0000-0001-5246-9351), Noemi Betancort (0000-0002-0156-3556), Ina Blümel, Christian Busse, Andreas Daniel, Felix Engel, Luca Ghiringhelli, Stephan Hachinger, Holger Israel (0000-0002-3045-4412), Naouel Karam, Norbert Kockmann, Angela Kranz, Richard Lenz, David Linke, Taras Petrenko, Lozana Rossenova (0000-0002-5190-1867), Daniela Schulz (0000-0003-3167-5089), Claus Weiland (0000-0003-0351-6523), Heinrich Widmann, Cord Wiljes (0000-0003-2528-5391), Mehtap Özaslan (0000-0001-8545-7576)

Version 1.0

2022-06-24

1. Motivation

In order to reach semantic interoperability and facilitate the implementation of the FAIR principles, several interrelated task areas have been identified that can benefit from cross-NFDI coordination. These include the harmonization and/or mapping of terminologies and the delivery of "FAIR" terminologies, to be used in connected domain-specific knowledge graphs, based on interoperable technical frameworks. Harmonization refers to finding an agreement on a common terminology, whereas mapping refers to an "agreement to disagree", while using a well-defined interface between terminologies.

Each consortium in the NFDI comes with their own set of terminologies, at different stages of maturity, and at different levels of FAIRness. Accordingly, ontologies and controlled vocabularies evolve in different time-frames and for different purposes. They are prone to have conceptual overlaps and - as in natural language - there are different ways of expressing the same concept. Mappings between equivalent representations of the same concept are thus essential for the federation of data that are structured with domain ontologies. The creation of formal, machine-readable mappings and their metadata needs to be guided by common standards. Additionally, the use of upper-level ontologies and the re-use of terms from existing terminologies need to be promoted. In addition to mapping (near) equivalent and semantically overlapping terms, disambiguation between terms "overloaded" with different meanings in different disciplines is another key task.

2. Objectives

- Increase the level of cross-domain interoperability within NFDI and to other infrastructures (EOSC) by provision of linkages and crosswalks between terminologies, ontologies and domain-specific knowledge graphs.
- Assessment of standards (i.e. formal, accessible, shared and broadly applicable language for knowledge representation) for the creation of mappings (e.g., SSSOM, OAEI), the formal documentation of mappings and their metadata in collaboration with the Terminology Services WG, as well as their (re-)evaluation and publication.
- Agree on an upper-level ontology to support broad semantic interoperability among domain-specific ontologies.
- Collect, amend, and harmonize quality standards for metadata annotations in the context of their practical application in NFDI and EOSC (cf. WG_Search_and_Harvesting.docx).
- Demonstration of how ontologies are applied to concrete metadata (cf. WG Cookbook(s), Guidance and Best Practices.docx).
- Reach out to and collaborate with communities of ontology maintainers to ensure that steps needed for harmonization are taken up into the original ontology specification as well as to promote the acceptance and implementation of changes to the source terminologies.
- Facilitating the FAIRification of terminologies used in the NFDI consortia by providing concrete examples and by embedding them in a FAIR "ecosystem" of services to reach semantic interoperability.

3. Work Plan: projected timeframes and milestones

Q3 – Q4 2022: Evaluating the state of the art
- Collect and assess ontologies used within the NFDI consortia (and beyond)
- Evaluate the use of a Crowd-Engineering approach (e.g., use the ORCID for contributions to a specific concept)
- Evaluate ways to get measurements on how often a service or a specific resource was used/requested.
- Review existing integration approaches such as KBpedia

Q1 2023: Overviews
- Create an overview of the state of terminologies and ontologies used in NFDI consortia and their respective domains.
- Create an overview of relevant activities within the NFDI consortia – as well as worldwide.

Q2 2023: Starting implementation, piloting and initializing the feedback process
- Coordinate cross-domain subgroups with representatives of each consortium who are already working on mappings and develop strategies for sharing skills and disseminating the results of each subgroup's cross-domain activities
- Provide a collaborative Git repository
- Coordinate testing sessions and feedback gathering with concrete cross-domain mappings and implementations in further services (e.g. KGs & TSs)

Q2 2023: Communication activities
- Give talks to working groups, NFDI consortias, and other groups of interests to get feedback
- Provide platforms for collaboration with ontology maintainer groups and other stakeholders
- Provide regular updates about WG activities
- Prepare for and participate in NFDI Conference 2023 and Section Workshops

Q3 – Q4 2023: Dissemination and documentation of outcomes
- Publication of a white paper summarizing the WG activities
- Talks, tutorials and workshops to working groups, NFDI consortia, and other groups of interests

Specific tasks:
- Initiate regular bi-monthly meetings of the whole working group
- Evaluate existing domain-agnostic vocabularies
- Evaluate existing upper-level ontologies
- Assessment of existing approaches to achieve "cross-domain knowledge sharing", e.g., minimum common metadata schemas
- Define minimum information standard for NFDI data publications required for semantic interoperability
- Identify and prioritize concrete needs for mappings across terminologies and/or knowledge graphs, by working groups staffed by all the involved consortia.
- Suggest recommendations on mapping tools, mapping creation, mapping validation, mapping documentation, and mapping publishing (e.g., on the Terminology Service)
- Provide example implementations via Terminology Service and research metadata repository applications that follow these newly set out mapping recommendations.

- Coordinate cross-domain subgroups which will work on defining and delivering specific mappings across ontologies and/or knowledge graphs that have been identified as important by individual consortia.

4. Initial Membership List
(At least 6 members from different institutions and at least 6 consortia)

Name	Consortium	Role in Workgroup	Email
Susanne Arndt	NFDI4Ing NFDI4Phys?	Contributor	susanne.arndt@tib.eu
Mehtap Özaslan	NFDI4Cat	Contributor	m.oezaslan@tu-braunschweig.de
Angela Kranz	DataPLANT	Contributor	a.kranz@fz-juelich.de
Noemi Betancort	KonsortSWD	Contributor	noemi.betancort@suub.uni-bremen.de
Holger Israel	NFDI4Phys et al.	Contributor or Co-Spoker	holger.israel@tib.eu
Naouel Karam	NFDI4Biodiversity	Contributor	karam@infai.org naouel.karam@fokus.fraunhofer.de
Roman Baum	NFDI4Health	Contributor	baum@zbmed.de
Claus Weiland	NFDI4Earth	Contributor	cweiland@senckenberg.de
Ina Blümel	NFDI4Culture	Contributor	ina.bluemel@tib.eu
Lozana Rossenova	NFDI4Culture	Contributor	lozana.rossenova@tib.eu
Felix Engel	NFDI4Ing	Contributor	felix.engel@tib.eu
Tobias Arera-Rütenik	NFDI4Objects	Contributor	tobias.arera-ruetenik@uni-bamberg.de
Christian Busse	NFDI4Immuno	Contributor	christian.busse@dkfz-heidelberg.de
Daniela Schulz	Text+	Contributor	schulz@hab.de
Taras Petrenko	NFDI4Cat	Contributor	taras.petrenko@hlrs.de
Cord Wiljes	NFDI-GS	Contributor	cord.wiljes@nfdi.de
Ivonne Anders	NFDI4Earth	Contributor	anders@dkrz.de

Heinrich Widmann	NFDI4Earth	Contributor	widmann@dkrz.de
Richard Lenz	NFDI4Cat	Contributor	richard.lenz@fau.de
David Linke	NFDI4Cat	Contributor	david.linke@catalysis.de
Luca Ghiringhelli	FAIRmat	Contributor	luca.ghiringhelli@physik.hu-berlin.de
Stephan Hachinger	NFDI4Earth, NFDI4Ing, etc.	Contributor	stephan.hachinger@lrz.de
Andreas Daniel	KonsortSWD	Contributor	daniel@dzhw.eu
Norbert Kockmann	NFDI4Cat	Contributor	norbert.kockmann@tu-dortmund.de

5. Adoption Plan

Once concrete needs for harmonization and/or mapping of terminologies have been identified, and subgroups tasked with deliberating the actual changes and amendments to ontologies and vocabularies have been set up, feedback into the consortia and the scientific communities they represent is going to be crucial for the success of the WG.

Given the NFDI's bottom-up approach, the consortia are the forums for discussion which terminologies should be used and how they should be modified and adopted. This WG is going to support the consortia in these decision processes, by providing expertise and feedback from the NFDI as a whole.

Adoption plans for individual cases of terminology development may differ, depending on the state of the art. They are going to be agreed on a case-by-case basis, guided by FAIR principles and Good Scientific Practice as overarching principle.

Represented Consortia

NFDI4Ing
NFDI4Phys
NFDI4Cat
DataPLANT
KonsortSWD
NFDI4Biodiversity
NFDI4Health
NFDI4Earth
NFDI4Culture
NFDI4Objects
NFDI4Immuno
Text+

Connected to other NFDI working groups

Search & Harvesting

Semantic Interoperability: Terminology Services
Cookbooks, Guidance, Best Practices
Knowledge Graphs
Working Group "Training Infrastructures" of Section Training & Education
Working Group "Data Integration" of section "Common Infrastructures"

Connected to other working groups outside NFDI
EOSC Task force Semantic Interoperability
RDA Data Granularity WG ?
W3C Web Ontology WG
The Ontology Matching Community and the Ontology Alignment Evaluation Initiative (OAEI)
The Simple Standard for Sharing Ontological Mappings (SSSOM)
The Open Biological and Biomedical Ontology Foundry (OBO)
Data for History Consortium
CIDOC Conceptual Reference Model Special Interest Group
LIDO Working Group
Project Anwendung Interoperabler Metadatenstandards (AIMS)
The FAIR Digital Objects Forum