



NFDI4Objects

Research Data Infrastructure
for the Material Remains of
Human History

TRAIL 4.2:

Implementing mapping processes for vocabularies related to site and object protection

Partner **Lead:** RGZM;

Co-applicants: RGZM, VZG, GDKE, DAI;

Participants: Hornemann Institute of the University of Applied Sciences und Arts (Weyer, Hentschel), University of Bamberg, Centre for Heritage Conservation Studies and Techniques (Arera-Rütenik), LVR-State Service for Archaeological Heritage (Claßen), The Role of Culture in Early Expansions of Humans - Project (C. Sommer);

External members: N4O Knowledge Modelling Cluster, N4O Cluster 3: Authority Files and community-driven Vocabularies Cluster, NFDI4Culture

Contact Florian Thiery M.Sc. / thiery@rgzm.de,

Dr. Allard Mees (RGZM) / mees@rgzm.de;

Dr. Tobias Arera-Rütenik (University of Bamberg) / tobias.arera-ruetenik@uni-bamberg.de

Summary

Different institutions working in domain-specific contexts document how sites and objects are conserved, preserved, restored and reconstructed in different ways. Therefore, different domain-specific authority files and community-driven vocabularies exist, are limited to the respective disciplines and reflect the specific logics of the activities. To promote synergies and linkages between these stakeholders, we need cross-disciplinary semantic alignment of these vocabularies. This addresses several phases of the research data cycle: capturing, inventorying, digitising, and providing, sharing and

integrating related (meta)data for research purposes. The aim of the TRAIL is to link terms in individual isolated systems by means of mapping relations in order to create a network of authority files and community-driven vocabularies from the fields of restoration and conservation, building research, architectural history and heritage management. This enables a multi-perspective overview of the biographies of (digitised) objects and makes it possible to relate objects in digital collections to each other. This will be done using two Software Application Services (SAS), DANTE and Cocoda. This will result in semantically aligned vocabularies, best practice white papers and OWL files defining the used relations.

Description

In this TRAIL, Authority File and Vocabulary Services (AVS) will be provided and used by N4O partners, and Interoperable Data Set Services (IntS) will be created. In addition, two existing Software Application Services (SAS) of VZG will be used: DANTE and Cocoda. The TRAIL will prototypically map various existing authority files and community-driven vocabularies to each other and to higher-level vocabularies. These vocabularies include Getty AAT/ULAN, Heritage Data (e.g. FISH), iDAI.world, the Archaeological Museum Hamburg, RGZM for restoration and conservation, and Bamberg for historical architecture: EwaGlos (Hornemann Institute) and ISCS-Stoneconservation (ICOMOS), Eagle (Europeana), Wortnetz Kultur, The Role of Culture in Early Expansions of Humans - project on Palaeolithic cultural phases and lithic vocabularies.

These authority files and community-driven vocabularies will (if they are not already) be integrated into DANTE as a shared vocabulary service. Through interdisciplinary dialogue, necessary properties beyond SKOS mapping relations¹ such as fuzziness and wobbliness² (see TRAIL 2.2) are worked out and semantically modelled in an OWL file. For the cross-vocabulary mapping, we will use VZG's Cocoda, which will be realised by the Hornemann Institut at the Hochschule für angewandte Wissenschaft und Kunst Hildesheim under the direction of Dr. Angela Weyer and Dipl.-Rest. Barbara Hentschel M.A. Regular evaluations in cooperation with TA6 will check the procedure and results of the mapping; we will check the significance of the defined properties with TAs 1, 2 and 3. The work will be documented within a best practice white paper. Since the measures relevant in TA4 can be understood as specific object events, semantic alignments (e.g. gold as a subclass of E57_Material³) of vocabulary terms to CIDOC-CRM classes will also be documented in a white paper and OWL file. The latter can be adopted directly in the conservation science related task T4.1.6. In addition, the semantic mapping of community-driven vocabularies relates directly to the interdisciplinary TRAIL 2.3. This TRAIL builds on the SKOS standard and lays the foundations for semantically extending its mapping relations (cf. TRAIL 2.2). The innovation in this TRAIL is that a domain-specific community generates semantic mappings across disciplines to develop proposals to better describe the various datasets, relate them to each other and evaluate them despite different subject-specific perspectives.

¹ <https://www.w3.org/TR/skos-reference/#mapping>, e.g. broadMatch, relatedMatch, closeMatch, exactMatch

² see also <https://github.com/zonination/perceptions>

³ <http://www.cidoc-crm.org/Entity/e57-material/version-6.2>

Relevance

The following aspect of the research data lifecycle is addressed: enrich and interpret. Data curators have semantically modelled authority files and community-driven vocabularies at their disposal to describe object protection measures. Infrastructure providers and system integrators can draw on a common vocabulary network. The alignment of vocabularies should facilitate data exchange between authorities, institutes and universities as well as continuous enrichment of the vocabularies.

The potential for other, non-participating communities is great, as interdisciplinary exchange is promoted through domain-specific mappings to general accepted vocabularies. Networking in other NFDI consortia, such as NFDI4Culture (TA2: Standards, data quality and curation; TA3: Research tools and data services), NFDI4Chem (Terminology Service⁴) and NFDI4Ing (SIG metadata & ontologies⁵) is justified by the interdisciplinary context of use of the targeted vocabularies (e.g. for materials, processes or buildings).

By integrating authority files and community-driven vocabularies into DANTE, the data are accessible through an open access protocol and they can be harvested and indexed. The main goal of this TRAIL is the semantic alignment of vocabularies to create qualified references to other data, which supports interoperability. W3C machine-readable community standards such as SKOS and RDF are used to ensure reusability.

For the NFDI, the development of a workflow for mapping authority files and community-driven vocabularies, incorporating the SKOS standard and its N4O-specific development, contributes to open science for RDM in the humanities and cultural sciences. The technology and methodology is also scalable and usable for other areas in NFDI with similar data structures (e.g. NFDICulture – [LOD.ACADEMY](https://www.lod.academy/)).

Deliverables

- Aligned authority files and community-driven vocabularies in Cocoda (**AVS**)
- Best practice white paper on the evaluation of SKOS properties incl. OWL file and blue paper (**IntS**)
- Best practice white paper on mapping vocabularies to CIDOC Classes incl. OWL file and blue paper (**IntS**)
- **N4O Commons**: white paper + blue paper + OWL files

⁴ <https://terminology.nfdi4chem.de/ts/index>

⁵ <https://nfdi4ing.de/community-hub/special-interest-groups-sig/metadata-ontologies/>

*FAIR*⁶ F4:RDA-F4-01M; A1.1:RDA-A1.1-01D; I1:RDA-I1-01D; R1.3:RDA-R1.3-02D

TRAILS related with TRAILS 4.1, 2.2, 2.3, and 2.7

⁶ Nach Tabelle 1 von Bahim, C., Casorrán-Amilburu, C., Dekkers, M., Herczog, E., Loozen, N., Repanas, K., ... Stall, S. (2020). The FAIR Data Maturity Model: An Approach to Harmonise FAIR Assessments. *Data Science Journal*, 19(1), 41. DOI: <http://doi.org/10.5334/dsj-2020-041> [cc by 4.0](#)