NFDI4Objects
Research Data Infrastructure for the Material Remains of Human History

TRAIL 2.5:

A workflow for enhancing iconography authority data in the Wikimedia Universe

Partner **Lead**: RGZM

Co-applicants: RGZM, SPK (Institute for Museum Research, SMB)

Participants: Technische Universität Braunschweig (Wienand), Heidelberg University (Börner), Goethe University Frankfurt (Tolle), University of Cologne (Eide), Berlin-Brandenburg Academy of Sciences and Humanities (Peter), Thüringisches Landesamt für Denkmalpflege und Archäologie; AdW Mainz (Deicke)

External members: N4O Linked Open Data Cluster, NFDI4Culture, Text+, Wikimedia Deutschland (Team Wissenschaft), German Maritime Museum (Bremerhaven)

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

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

Prof. Dr. Johannes Wienand (Technische Universität Braunschweig) / j.wienand@tu-bs.de,

Dr. Susanne Börner (ZAN, Heidelberg University) / susanne.boerner@zaw.uni-heidelberg.de

Summary

Contextualising heterogeneous object data with regard to their geographical, temporal and iconographic properties is a great challenge. In this TRAIL, this challenge will be tackled using numismatics as an example, since standardisation for describing objects in this domain is fairly advanced. Although in the field of iconography, machine-

readable term catalogues and identifiers (URI) are already available digitally (e.g. IconClass, Getty AAT), these are too unspecific for numismatic research questions. Additionally, descriptive elements such as texts and images are necessary for a complete classification and contextualisation of the imagery of objects. To improve this situation, selected community-driven vocabularies will be made available via the DANTE vocabulary server and enriched with further descriptive elements such as free texts (incl. references) and media data published via the Wikimedia Universe. This TRAIL addresses the describe, publish and qualify aspects of the research data lifecycle. It will produce a model workflow as a best practice white paper describing the linking of iconography entries in the DANTE vocabulary server with resources in the *Wikimedia Universe*. For this procedure, supplementary OER are also created and used in university tandem courses.

Description

The subject of this TRAIL is iconography data as part of coin descriptions for which community-driven vocabularies are already publicly available. Some are already accessible via the VZG's vocabulary server, DANTE, others still need to be integrated into this service. The specific challenge of this TRAIL is the formulation of community-approved and consistent explanations of vocabulary terms across different originating systems with the help of free texts and images, which are either already available in free community hubs such as the *Wikimedia Universe* (e.g. Wikidata, Wikimedia Commons, Wikipedia) or can be uploaded there.

In this TRAIL, selected Authority File and Vocabulary Services (AVS) provided by N4O partners are used and Data Services (DaS) are created:

- Iconography Thesaurus on Greek Numismatics (Berlin-Brandenburg Academy of Sciences and Humanities)
- NUMiD / Greek Coinage Typology Index (Technische Universität Braunschweig)
- Roman coins with depictions of ships (NAVIS3) (RGZM)

The aim of the TRAIL is to formulate a model workflow that both enables quality assurance of the subject-specific concepts and resources available in the Wikimedia Universe and ensures the link between the vocabulary terms in DANTE and the resources in Wikimedia. This is being carried out under the leadership of Prof. Dr Johannes Wienand (Technische Universität Braunschweig; NUMiD network).

A solution approach to meet these challenges will be implemented as a multi-level workflow that supplements the description options of the Simple Knowledge Organisation System (SKOS, e.g. semantic relationships, labels, documentary notes) with elements such as free texts (incl. references) and images:

- Transformation of community-driven vocabularies according to SKOS

- Publication of these SKOSified community-driven vocabularies via the DANTE vocabulary server

- Integration of extended data on vocabulary terms through new or existing resources in the *Wikimedia Universe*: Wikidata (as Linked Open Data Community Hub) – Wikipedia (long description) - Wikimedia Commons (e.g. images, documents)

- Creation of bidirectional links between DANTE and Wikidata

The resources created in this way can be dynamically integrated into IT architectures and domain-specific research collections (e.g. NAVIS) for contextualisation.

To convey the workflow, accompanying OERs will be created, which will be tested and used in project seminars at the Heidelberg University (Faculty of Philosophy, MA degree programme Cultural Heritage and the Protection of Cultural Property). Tandem events with numismatists and developers are planned there. The OERs created will be processed and integrated as a reference implementation into the NumiScience.de e-learning platform developed there by the Numismatic Association in Baden-Württemberg (NV BW). In addition, the OER can be used in any course via the Qualification Services (QaS) offered in TA6.

The workflow will result in *AVS*published via DANTE*,* DaS with integrated DiS in the Wikimedia Universe and QuaS via the e-learning platform of NV BW. The SKOS standard of W3C is used to model the AVS. In the *Wikimedia Universe*, the Wikidata model can be applied as a community standard.

Thanks to the enrichment and bidirectionally linkage of vocabulary terms and concepts the basis for publishing the object data they describe in the LOD Cloud is created and the (semantic) searchability with available tools is enabled.

Relevance

In this TRAIL, various community-driven vocabularies on iconography are enriched and their terms defined using the Wikimedia Universe with texts (Wikipedia), images (Wikimedia Commons) and semantic metadata (Wikidata) published under free Creative Commons licences.

Even though the workflow is modelled on iconography examples, it is transferable to other types of metadata used to describe objects. Due to its reusability in other subject domains, e.g. conservation science (TA4), material analyses (TA3) or construction research (TA1 and TA4), different users and stakeholders can benefit from the results, including scientists and data curators, university teachers, students, citizen scientists. Citizen science is an area with particularly high potential, as the data are created in the Wikimedia Universe with the support of Wikimedia Deutschland (Team Wissenschaft), enabling participation and access for communities beyond N4O. Synergies can be forged with other humanities and culture-related NFDI consortia – NFDI4Culture (TA2: Standards, data quality and curation; TA3: FAIRification tools and data services; TA4:

Data publication and data availability), Text+ (Cross-cutting topic data, service and software quality frameworks); NFDI4Memory (TA2: Data connectivity; TA3: Data services). In addition, the workflow can be connected to those parts of the linked data community, which focus on humanities and cultural data and which are primarily present in associations like Linked Pasts, CAA International and CAA Germany.

All four elements of FAIR are addressed in this TRAIL: the use of community-driven vocabularies, specified by qualified references to external resources and concepts, significantly contributes to the interoperability of descriptive object data. Publishing these enriched terms and concepts in DANTE and the *Wikimedia Universe* makes them especially findable and accessible. Furthermore, the use of machine-readable community standards such as SKOS, RDF and the Wikidata data model ensures extended reusability.

For the NFDI, developing a workflow involving standards and *Wikimedia Universe* is a technically and methodologically relevant contribution to Open Science in iconography research in the humanities and cultural sciences. The digital methods used are scalable and usable for other NFDI consortia and working groups with similar data structures (e.g. the AG Archäologiethesaurus, working on a common vocabulary for the cultural heritage authorities, and Arbeitskreis Digitale Kunstgeschichte, the working group on digital art history).

Deliverables

- Community-driven vocabularies on iconography in DANTE (**AVS**)
- Iconography data in the Wikimedia Universe (**DaS**)
- Workflow described in a white paper and disseminated via the e-learning platform NumiScience.de and via courses at Heidelberg University (**QuaS**)
- **N4O Commons**: white paper / OER

Work plan

- Year 2, Month 3: Community-driven vocabularies integrated in DANTE
- Year 2, Month 6: Workflow creation completed
- Year 2, Month 9: Resources integrated in the Wikimedia Universe (Wikidata, Wikimedia Commons, Wikipedia)
- Year 2, Month 12: OER *materials*, courses held / workflow white paper completed

FAIR[1]	F4:RDA-F4-01M; A1.1:RDA-A1.1-01D; I1:RDA-I1-01D; R1.3:RDA-R1.3-02D
TRAILS	based on TRAIL 2.1

[1] 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