TRAIL 2.7:

An URI lookup and resolve service for LOD resources related to collection research

Partner   **Lead**: RGZM;

Co-applicants: RGZM, VZG, DAI, SPK (Institute for Museum Research);

Participants: Goethe University Frankfurt (Tolle), University of Cologne (Eide), AdW Mainz (A. Deicke);

External members: N4O Knowledge Modelling Cluster, CAA International - SIG on Semantics and LOUD in Archaeology, Linked Pasts, NFDI4Culture, NFDI4Memory, NFDI4Earth, mainzed

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

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

Prof. Dr. Øyvind Eide (University of Cologne) / oeide@uni-koeln.de

Summary

Research domain-specific modelled Linked Open Data (LOD), e.g. distributed thesauri, must be dynamically integrated into collection research databases to enable common understanding of the research data. The registry entries recorded in the *Catalogue of authority files and community-driven vocabularies Service* and the associated standards and ontologies elicited form the starting point for a much-needed FAIRification tool for LOD in archaeology: the re3dragon tool (REsearch REsource REgistry for Data Dragons). The re3dragon makes it possible to collect both overarching authority files and community-driven vocabularies, as well as LOD published from domain-specific research collections, to search in them and to yield individual terms in a machine-readable,

standardised data format. The TRAIL therefore addresses the enrich & interpret and share, publish & discover aspects of the research data lifecycle and contributes to the FAIRification of domain-specific concepts. The aim of this TRAIL is, first, to enhance the functionality of the tool by integrating in the search routines new authority files, community-driven vocabularies and LOD resources based on the white paper derived from TRAIL 2.1 and name authority files from TRAIL 2.4. Second, we will expand the technical specification of the JSKOS data exchange format, which is already supported by re3dragon, to include missing object-specific criteria.

Description

Scholars and scientists need authority files and community-driven vocabularies that describe research data with qualifying references to enable common understanding of research data. Using such descriptive data in individual research requires FAIRification tools that support the search and the use of suitable terms and concepts to describe objects. The TRAIL will enhance a Software Application Service (SAS): re3dragon, developed by the RGZM.[1] The re3dragon is a JAVA-based FAIRification tool enabling a data reconciliation service that can be used to collect and verify authority files and community-driven vocabularies (e.g. GND, iDAI.chronontology, ISIL) as well as published, specialised LOD (e.g. approximated kiln site regions from Samian[2] or amphora types from Roman Open Data[3]), to search them dynamically, to resolve individual terms and to output the results in a standardised format. The re3dragon tool is integrated into the RGZM's RDM infrastructure and made available to other interested parties via GitHub, uses the JSKOS format implemented by the VZG,[4] and works closely with the CAA SIG on Semantics and LOUD in Archaeology[5] and Linked Pasts[6]. The TRAIL will include Authority File and Vocabulary Services (AVS) and Data Services (DaS) provided by N4O partners (e.g. existing vocabulary servers with APIs such as DANTE) as well as by other, including international actors (e.g. Getty AAT, Iconclass, GND). The data to be included in this TRAIL will be evaluated beforehand in TRAILs 2.1 and 2.4 and coordinated with TA6 and M6.1. Published and machine-readable authority files and community-driven vocabularies are often semantically modelled differently. Some are based on the W3C standard SKOS.[7] The VZG has taken first steps in the standardisation process for a JSON exchange format: JSKOS. However, the existing version of JSKOS is the basis for a standardised exchange of LOD resources with the help of re3dragon. In the process of integrating further resources, the JSKOS format will be further developed in this TRAIL in cooperation with the VZG using the results of TRAIL 2.2 in an object-specific way. The challenge for this TRAIL is to balance standardised responses to: on the one hand differently structured LOD resources and on the other hand results of searches on decentralised, heterogeneous vocabularies from different research domains, institutions and

---

[1] c.f. https://github.com/RGZM/re3dragon; Thiery, F 2021 *re3dragon - REsearch REsource REgistry for DataDragons*. Scientific IT Research Tools. Mainz: Römisch-Germanisches Zentralmuseum. DOI: 10.5281/zenodo.5338740.
[2] c.f. https://rgzm.github.io/samian-lod/datapub/; Thiery, F, Mees, A and Gottwald, D 2020 *Linked Open Samian Ware*. Scientific IT LOD. Mainz: Römisch-Germanisches Zentralmuseum. DOI: 10.5281/zenodo.4305708.
[3] c.f. https://romanopendata.eu
[4] c.f. https://gbv.github.io/jskos/
[5] c.f. https://caa-international.org/special-interest-groups/; http://datadragon.link
[6] c.f. https://linkedpasts.hcommons.org
[7] c.f. https://www.w3.org/TR/2009/NOTE-skos-primer-20090818/

granularities. Prof. Dr Øyvind Eide and his team at the University of Cologne will model and implement this. At present, these resources are individually integrated into collection research databases and implemented in individual solutions. The extended re3dragon tool offers an innovative single-access solution that can dynamically integrate a larger number of LOD resources into research collections.

Relevance

The TRAIL addresses the enrich & interpret and share, publish & discover aspects of the research data lifecycle. Users and stakeholders who will benefit most are scientists, data curators and RSEs. The FAIRification tool provides a common machine-readable interface to look up and resolve authority files, community-driven vocabularies and other LOD resources. Stakeholders can search distributed vocabularies, use the results to enrich datasets with FAIR-compliant vocabularies and develop individual solutions in database systems for collection research[8]. The potential for the other communities involved in N4O is great, as functionality enhancement of the re3dragon tool and object-specific adaptation of JSKOS is useful for all disciplines. Networking in other NFDI consortia, such as NFDI4Culture (TA2: Standards, data quality and curation; TA3: FAIRification tools and data services), NFDI4Memory (TA2 Data connectivity; TA3 Data services) or NFDI4Earth (FAIR Data & Interoperability and Research Software interest groups) is justified by the interdisciplinary context in which the targeted vocabularies are used. The overall aim of the TRAIL is to enable object databases to use FAIR vocabularies in their metadata. Collecting LOD resources[9], authority files and community-driven vocabularies generate enriched metadata that increases discoverability of the accompanying object data. The API standardises knowledge representation, which contributes to the interoperability of data; the free access protocol unifies access to it. W3C machine-readable community standards such as SKOS and RDF are used to ensure reusability. For the NFDI as a whole, re3dragon ensures a technically and methodologically relevant contribution by providing a discovery tool for LOD resources. In this TRAIL, especially TA2 specific resources are used (e.g. white paper from TRAIL 2.1 and name authority files from TRAIL 2.4), but the technology and methodology is scalable and reusable for other areas in NFDI with similar data structures (e.g. NFDI4Culture: LOD.ACADEMY[10]).

Deliverables

- Development/implementation of the FAIRification tool re3dragon (**SAS**)
- Further development of the JSKOS data format (**IntS**)
- **N4O Commons**: White Paper / Blue Paper / community standard

---

[8] c.f. https://commons.wikimedia.org/wiki/File:Collection_Research_Network.png
[9] c.f. https://github.com/RGZM/re3dragon/blob/master/dragonlairs.ttl
[10] c.f. https://lod.academy/site/

FAIR[11]   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, TRAIL 2.4, TRAIL 2.2

---

[11] 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