



NFDI4Objects

Research Data Infrastructure
for the Material Remains of
Human History

TRAIL 2.4:

Digital provenance research: methodologies, concepts and name authority files

Partner **Lead:** SPK

Co-applicants: SPK (SMB Zentralarchiv, Institute for Museum Research, Justiar-
iat), RGZM, VZG

Participants: Technische Universität Berlin (Prof. M. Hopp), DNB (GND), NUMID
Verbund, Heidelberg University Library (Maria Effinger; Propylaeum-Vitae)

External members: Arbeitskreis Provenienzforschung e. V. (AG Digitale Proveni-
enzforschung), Deutscher Museumsbund, Network Colonial Contexts - EVIFA,
NFDI4Culture, N4O Linked Open Data Cluster, N4O Authority Data Cluster, Wiki-
media Deutschland (Team Wissenschaft)

Contact Prof. Dr. Bernhard Weisser / b.weisser@smb.spk-berlin.de;

Dr. Petra Winter / p.winter@smb.spk-berlin.de;

Felix F. Schäfer (SPK) / f.schaefer@hv.spk-berlin.de;

Prof. Dr. Meike Hopp (TU Berlin) / meike.hopp@tu-berlin.de

Summary

The provenance or origin of an object is central to its biography. The TRAIL deals with the methodological foundations and data-related requirements for researching the previous owners and sellers of objects which are now in public ownership. It addresses the describe, qualify and publish aspects of the research data lifecycle. The aim of the TRAIL is to categorise and catalogue existing resources and create a new paradigm for name authority files (TU Berlin, Prof. Dr. Meike Hopp), which will be published using

the DANTE vocabulary server and, if possible, international authority file repositories. This will prioritise enriched and qualified (meta)data by relevance. Only shared name authority files will allow the investigation of networks, mechanisms and interests of previous owners or sellers of objects and to link these to further sources. The semantic network generated in this way, together with specific data models (including aspects of vagueness and uncertainties, see TRAIL 2.2), is a prerequisite for other scientists to analyse the initial data in more depth. A white paper will summarise the relevance and potential of name authority files and provide methodology to enrich them in other data collections.

Description

The TRAIL contributes to improving data quality in various Authority File and Vocabulary Services (AVS), Interoperability Services (IntS) and Data Services (DaS), based on the person related metadata on objects and their provenance. To achieve this, existing representative database inventories from several projects can be meaningfully merged and combined in terms of content and technology. This includes ongoing and completed research on the art trade during National Socialism (M. Hopp, TU Berlin) and on university collections (NUMiD) as well as the BKM-funded project on digital provenance research in numismatic collections of the period, “Projekt zur digitalen Provenienzforschung im Münzkabinett: Veräußerer und Vorbesitzer zwischen 1933 und 1945” (SPK).

Provenances of objects in public collections can be very diverse, as they come into museum institutions through purchase, transfer, exchange or donation. Knowledge about previous owners, sellers and transfer of custody are highly relevant for the interpretive histories and biographies of the objects. Persons who are important for the history of objects deserve name authority files, as it is already the daily practice for authors of scientific publications. This is the only way to gain insights into interrelated object groups and networks across collections.

International (VIAF, LCCN, ULAN) and national (GND) authority file services will be used to create new name authority files. Propylaeum.Vitae, the Getty Provenance Index® (GPI) and PROVEANA can also be used, although the latter is still very incomplete, particularly regarding standardisation of core metadata elements, and extremely incoherent in terms of indexing (occupation titles, etc.). The greatest challenge is that information about persons comes from very heterogeneous source material and can be incomplete, incorrect, contradictory or ambiguous in itself (e.g. missing first names or place references; incorrect spelling of proper names). Thus it is often impossible to clearly identify the named persons. This presents a great opportunity for the TRAIL, because creating interconnected entities (in the medium term also by using named entity recognition) and identifying connections (e.g. through object convolutes, collections or find contexts) can bring together incomplete personal information scattered across different source holdings. This creates the basis for valid authority files of mostly historical persons.

Up to now, provenance researchers have often focused on qualitative analyses of individual cases published in dossiers. Here, however, metadata is generated extensively on the basis of many objects, which makes it possible to link external and internal resources quantitatively, semantically and interoperably. Incomplete or fuzzy data should not be rejected as authority files, but data should be represented in a way that explicitly depicts these uncertainties. Through reasoning tools, unknown network links can be revealed, which then can help to enrich other datasets.

Relevance

In line with TA2, the TRAIL is a case study in enriching data on objects from collection contexts during the process of indexing (collect, assure, describe). By improving the quality of the metadata, e.g. by linking with newly created authority files, object data is easier to discover, integrate and analyse.

Researchers, data curators and university teachers will benefit from the results in two ways, scientifically and methodologically. First, high-quality name authority files will be created and used in model datasets; second, a white paper will discuss challenges, workflows and services for the creation, quality assurance and provision of object-relevant authority files. Digital provenance research and its targeted support through NFDI services will be addressed in two consecutive seminars at TU Berlin (M. Hopp) and HU Berlin (B. Weisser), which will focus on the subject-specific relevance and practical handling of valid name authority files. Participants will share their findings in presentations, essays and a public student colloquium. Further, the results are disseminated to the academic community via the Arbeitskreis Provenienzforschung e. V., whose spokesperson is Prof. Meike Hopp.

The TRAIL will also identify possibilities and challenges of working with object-related name authority files provided by the Deutsche Nationalbibliothek and the GND. Via Prof. Hopp and the Prussian Cultural Heritage Foundation, there are numerous personal and thematic overlaps with NFDI4culture. Even though this consortium focuses on other cultural assets and object categories, the data challenges are comparable. Via common name authority files cultural objects can also be related with objects from science collections. One measure of success with object-related name authority files is the extent to which persons become visible not only as authors of (scholarly) texts, but also as influential figures within object biographies.

In this TRAIL, all FAIR principles are addressed in practice on different levels. Persistent identifiers are used for the data; all authority files that are created are published with traceable source information; where possible, internationally established authority file repositories such as VIAF are used, then national ones such as GND and finally subject-specific ones such as the authority file portal of the SPK Münzkabinett; data is exchanged via standard protocols. As the data is integrated into VZG's DANTE system, its ongoing curation and improvement is ensured. This institutional connection between the SPK and VZG ensures maximum long-term availability of digitised material independent of the individual programme.

Digital provenance research is key to contextualising objects in terms of their provenance, authenticity and interpretation. Originally existing contexts of use, expropriation or relocation can be reconstructed, as can different motivations and mechanisms of (illegitimate) trade, acquisition and storage of cultural objects. This TRAIL not only illustrates the importance of metadata and concepts (particularly name authority files and LODs) for the recording of objects and the enrichment of (meta-)data. It also initiates exchange between curators of authority files at the GND and the specialist community engaged with object collections.

Deliverables

- White paper on the inventory and evaluation of relevant datasets and services as well as on the object-specific management of (name) authority files
- New and/or enriched name authority files for provenance research of archaeological and ethnological objects (**AVS**)
- N4O Commons: White paper

Work plan

- Year 1, Month 1–6: Focus I: Provenance data on archaeology objects and their legal basis
- Year 1, Month 1–12: Focus II: Provenance data on ethnology objects and their legal basis
- Year 1, Month 1, 4, 9, 11. Workshops (e.g. on dealing with ambiguous personal data and on the role of the GND with the DNB); the results of the workshops will be included in the final white paper
- Year 1, Month 12: Completion of white paper

FAIR¹ F4:RDA-F4-01M; A1.1:RDA-A1.1-01D; I1:RDA-I1-01D; R1.3:RDA-R1.3-02D

TRAILS part of TRAIL 2.1, related with TRAIL 2.3, basis for TRAIL 2.7

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