NFDI4Objects
Research Data Infrastructure for the Material Remains of Human History

TRAIL 2.2:

Evaluation of fuzziness and wobbliness in numismatics and ceramology

Partner **Lead**: RGZM

Co-applicants: RGZM, VZG, DAI (Römisch-Germanische-Komission), GDKE, SPK

Participants: Goethe University Frankfurt (Tolle), University of Cologne (Eide), CAA Germany, Museum für Naturkunde Berlin (Bölling), University of Bamberg (Arera-Rütenik) ADW Mainz (A. Deicke)

External members: N4O Cluster Knowledge Modelling, CAA International - SIG on Semantics and LOUD in Archaeology, NFDI4Culture, NFDI4Earth, mainzed (Mara)

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

Florian Thiery M.Sc. (RGZM) / thiery@rgzm.de,

Dr Karsten Tolle (Frankfurt Big Data Lab, Goethe University Frankfurt) / K.Tolle@em.uni-frankfurt.de,

David Wigg-Wolf (Römisch-Germanische-Komission) / david.wigg-wolf@dainst.de

Summary

Both exact and interpretive statements play a central role in the exploration of objects. The latter can be strongly influenced by authors' hidden assumptions, incomplete object preservation or imprecise technical concepts. Uniform modelling of the associated fuzziness and wobbliness in research data is a challenge, as different approaches are developed or implemented in research domains. This TRAIL addresses the enrich &

interpret and reuse & cite phases of the research data lifecycle. The aim is to collect, evaluate and extend existing modelling approaches to fuzziness and wobbliness (e.g. uncertainty, vagueness, accuracy and precision) in research data in a case study of collection-related numismatic and ceramic research. This will identify common intersections that enable transformations from one research discipline-specific modelling approach to another. Besides collecting mathematical, IT and other relevant implications of existing modelling approaches, this TRAIL will implement prototype modelling approaches in use cases and embed them in university courses, enabling scientists and students to evaluate the advantages and disadvantages in a verifiable way. White and blue papers outlining best practices and transformation rules for modelling concepts will be published so that the solutions found for numismatic and ceramic research can be transferred and adapted to other research domains.

Description

NFDI4Culture generally postulates that: "Metadata describing material and immaterial cultural assets itself can serve as research data. For this reason, metadata in the cultural heritage domain must allow the mapping of blurriness and uncertainty. Complementary or contradictory information also shall be captured in a human and machine-readable way."[1] Various modelling approaches to fuzziness and wobbliness in data about objects exist in different contexts (e.g. interpretations in the humanities and cultural sciences, experiments in the natural sciences, measurements in the engineering sciences). These are independent of each other and are not yet compatible. As a result, scholars and data curators cannot express the immanent fuzziness, wobbliness, accuracy and precision of relevant statements about objects (e.g. iconographic or typological attributions, cultural-historical dating, relative geographical indications) in a uniform way. The lack of comparability and explicitness in fuzzy and wobbly data leads to a lack of transparency with regard to data quality, obscures hidden assumptions and prevents processing (e.g. when comparing data of objects consisting of different materials).

The challenge lies in the different methods for deriving a statement about an object; for instance, a coin or a ceramic vessel are dated based on different criteria. The statements each contain individual dispositions of implicit fuzziness and wobbliness, which come into play when two objects from different find categories are to be related to each other. It does not seem feasible to solve the variety of domain-specific approaches with a single standard model for mapping fuzziness and wobbliness. Instead, existing modelling approaches (e.g. Tolle/Wigg-Wolf, 2015;[2] Unold et al., 2019;[3] Bruhn et al.,

[1] R. Altenhöner, I. Blümel, F. Boehm, J. Bove, K. Bicher, C. Bracht, O. Brand, L. Dieckmann, M. Effinger, M. Hagener, A. Hammes, L. Heller, A. Kailus, H. Kohle, J. Ludwig, A. Münzmay, S. Pittroff, M. Razum, D. Röwenstrunk, H. Sack, H. Simon, D. Schmidt, T. Schrade, A.-V. Walzel, B. Wiermann (2020): *NFDI4Culture - Consortium for research data on material and immaterial cultural heritage*. Research Ideas and Outcomes 6: e57036. DOI: 10.3897/rio.6.e57036.

[2] K. Tolle / D. Wigg-Wolf (2015): Uncertainty Handling for Ancient Coinage. CAA2014. 21st Century Archaeology. Concepts, methods and tools. Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology, pp. 171-178.

[3] M. Unold / F. Thiery / A. Mees (2019): *Academic Meta Tool. Ein Web-Tool zur Modellierung von Vagheit*. ZfdG - Zeitschrift für digitale Geisteswissenschaften, Sonderband 4. DOI: 10.17175/sb004_004.

2015;[4] Holmen et al., 2010[5]) should be further developed based on research domain-specific best practices and community standards to increase mutual compatibility. Furthermore, transformation rules are needed that enable semantic and technical compatibility between the modelling approaches. One way of doing this is to implement graph-based LOD models by adding quadruples with blank nodes or inserting shortcuts through extra properties.[6]

As the basis for creating an Interoperable Data Set Service (IntS) the following existing concepts are included in the TRAIL.

- CIDOC CRM; CRM extensions: CRMgeo, CRMarchaeo, CRMsci, CRMtex, CRMinf; LIDO-terminology; Nomisma; SKOS; W3C Uncertainty Ontology; W3C Time Ontology; PeriodO; Pleiades; Getty Vocabularies; Linked Places (Linked Pasts); Nomisma

During the TRAIL, data from numismatic and ceramic research will be used to develop best practices, as experts in these domains have been discussing these questions for a long time and can provide a good basis for testing new approaches. The following Data Services (DaS) and Software Application Services (SAS) of the N4O consortium are included:

- Academic Meta Tool (mainzed, RGZM); Samian Research, NAVIS (RGZM); Dating Mechanism Project / Alligator Tool (RGZM); Antike Fundmünzen in Europa-Römisch-Germanische-Komission (Römisch-Germanische-Komission/Uni Frankfurt); iDAI.gazetteer (DAI); iDAI.ChronOntology (DAI, i3mainz); Cocoda/coli-conc (VZG)

The TRAIL will produce white and blue papers that document the enhanced modelling approaches to fuzziness and wobbliness, as well as the transformation rules, in a formally structured way. Furthermore, a reference implementation will be set up to test the practicability and portability of the solutions to thematically closely related problems. The following databases on collection objects are used for this purpose:

- NAVIS3 database (RGZM): Fuzziness and wobbliness regarding ambiguous assignments in ship depictions on Roman coins[7]
- SFB-Projekt 933 Materiale Textkulturen, A11: Graffiti auf römischen Goldmünzen (Heidelberg University): Fuzziness and wobbliness regarding presumed reading directions and spatial locations of graffiti on Roman gold coins

[4] K.-C. Bruhn / T. Engel / T. Kohr / D. Gronenborn (2015): *Integrating Complex Archaeological Datasets from the Neolithic in a Web-Based GIS*. CAA2014. 21st Century Archaeology. Concepts, methods and tools. Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology, pp. 341-348.
[5] J. Holmen / C.-E. Ore (2010): *Deducing event chronology in a cultural heritage documentation system* in Making History Interactive. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference, Ed.J. W. Crawford and D. Koller Oxford: Archaeopress, 2010.
[6] F. Thiery / A. Mees / K. Tolle / D. Wigg-Wolf (2021): *Nomisma Uncertainty and AMT Vagueness Modelling Approaches*. Zenodo. DOI: 10.5281/zenodo.5520977.
[7] F. Thiery / A. Mees (2018): *Taming Ambiguity - Dealing with doubts in archaeological datasets using LOD*, CAA Tübingen. DOI: 10.5281/zenodo.1200111.

- Samian Research (RGZM): Fuzziness and wobbliness concerning intended potter's attributions, e.g. vessel makers cursive signatures inscribed before kiln firing[8]

In addition, suitable OER for the N4O Qualification Service (QuaS) are being developed in cooperation with TA6 as part of courses at the Universität Köln (Faculty of Philosophy) with the support of teachers (e.g. Prof. Øyvind Eide) who have experience in modelling fuzzy data in the humanities and cultural sciences. Master's students in Information Processing, Media Informatics and Archaeoinformatics will work on implementations of the technical parts of the white and blue papers, and discuss in a report their suggestions for improvement and further developments of the applied methods.

By formulating transformation rules, this TRAIL for the first time establishes comparability and explicitness of modelling approaches to fuzziness and wobbliness. This contributes to improving data quality, makes hidden assumptions more transparent and object data more interoperable in collection research (e.g. through reasoning or AI technologies[9]).

Relevance

The TRAIL addresses the enrich & interpret and reuse & cite phases of the research data lifecycle. Users and stakeholders who will benefit most are scholars and data curators. The approaches developed about fuzziness and wobbliness, and the deduced transformation rules, will provide a methodological basis for measures to enrich research data (cf. TRAIL 2.6).

This has great potential for other communities, as semantically explicit, structured, modelling of statements on fuzziness and wobbliness and their machine-readable representation is needed in other scientific domains of the NFDI to make the quality of data transparent and comprehensible. In individual disciplines related to collection research, this topic is already being addressed and actively discussed (e.g. CAA International, especially CAA SIG Data-Dragon). Master's students in various digital humanities curricula will also benefit from this TRAIL through participation in courses.

It will result in collaboration with other NFDI consortia, specifically NFDI4Culture (TA2: Standards, data quality and curation, TA3: Research tools and data services); NFDI4Memory (TA1 Data quality, TA2 Data connectivity); NFDI4Earth (FAIR Data & Interoperability interest group), to adapt and develop modelling strategies.

Two elements of FAIR are addressed in this TRAIL: interoperability and reusability. This TRAIL will create the basis for precise and explicit descriptions of a large number of relevant object properties, which can be reused granularly. The application of the LOD principles[10] (i.e. machine-readable knowledge representations expressed in standardised formats as well as qualified references to other data) and the W3C standard

[8] F. Thiery / A. Mees (2021): *Linked Open Samian Ware - AMT Perspective*, Zenodo. DOI: 10.5281/zenodo.5415571
[9] e.g. K. Tolle (2021): *Pattern recognition on ancient coins*. https://youtu.be/QKiPuWLcsBo.
[10] cf. https://www.w3.org/DesignIssues/LinkedData.html

Resource Description Framework (RDF) in a community-specific version ensures that the data is machine readable and interoperable.

For the NFDI, the white and blue papers, and transformation rules, make a technical and methodological contribution to the traceability and thus reusability of fuzzy and wobbly statements in object-related collection data. The digital methods used are scalable and reusable for other NFDI consortia and working groups with similar data structures (e.g. NFDI4Culture, working group "Graphen & Netzwerke"[11] of the "Verband Digital Humanities im deutschsprachigen Raum", as well as the Text Encoding Initiative[12] TEI-ONTO-SIG).

Deliverables

- Develop community standards and data-driven approaches to fuzzy wobbling semantic modelling (IntS, Best-Practice White Paper, Blue Paper)
- Best practices for modelling fuzzy wobbling semantic relations and transformation rules between modelling concepts (IntS, Best-Practice White Paper, Blue Paper)
- OERs for communicating the results in teaching (QuaS)
- N4O Commons: White Paper, Blue Paper

FAIR[13]	I1:RDA-I1-01D; R1.3:RDA-R1.3-02D
TRAILS	part of TRAIL 2.1, related with TRAIL 2.3, basis for TRAILs 2.6 and 2.7

[11] e.g. https://graphentechnologien.hypotheses.org/tagungen/graphentechnologien-2018
[12] e.g. certainty element https://www.tei-c.org/release/doc/tei-p5-doc/de/html/ref-certainty.html
[13] 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