Harmonizing and Using Numismatic Linked Data in Digital Humanities Research and Application Development: Case DigiNUMA

Heikki Rantala[1][0000−0002−4716−6564]*, Eljas Oksanen[2,1][0000−0002−7468−9256], and Eero Hyvönen[1,2][0000−0003−1695−5840]

[1] Semantic Computing Research Group (SeCo)
Aalto University and University of Helsinki, Finland
https://seco.cs.aalto.fi, firstname.lastname@aalto.fi
[2] HELDIG – Helsinki Centre for Digital Humanities, University of Helsinki, Finland
https://heldig.fi, firstname.lastname@helsinki.fi

Abstract. This paper outlines the ongoing work of the DigiNUMA project for creating solutions in data harmonisation, analysis, and dissemination of pan-European archaeological and numismatic Cultural Heritage, using linked data and semantic web technologies. The project focuses on Viking Age (800–1150 AD) Finnish and English numismatic data as a case study. A broader context is gained by research into harmonizing collection data of the National Museum of Finland, the British Museum, and the Fitzwilliam Museum in Cambridege for compatibility with the international Nomisma.org ontology, and by creating tools that can be used to work with other Nomisma.org datasets.

1 Introduction and Related Work

During the recent years the number of archaeological finds made by the public across Europe, mainly through hobby metal-detecting, has grown considerably. Coins form a special case of finds worth concentrating on for several reasons. Coin finds are relatively easy to recognise when found in the ground with a detector and are usually the most numerous object type reported by the public.[5] Coin types can be identified more precisely than other common finds, producing higher quality record data and making them specially suitable for Digital Humanities (DH) analysis; for example, dates and places of minting can often be determined with reference to existing numismatic scholarship. Historical coins also moved internationally, making harmonizing and comparing international data especially relevant.

DigiNUMA – Digital Solutions for European Numismatic Heritage[3] [6] is an ongoing research project that responds to a set of new needs in Cultural Heritage (CH) data management, research, and dissemination using Linked Open Data (LOD). The project collaborates with two international LOD data projects in

[3] Project homepage: https://seco.cs.aalto.fi/projects/diginuma

archaeological CH: ARIADNEplus[4] [8] and Nomisma.org[5] [2]. ARIADNEplus is a pan-European research infrastructure and aggregation project for all archaeological data, while Nomisma.org concerns numismatic data.

The project contributes to the state-of-the-art by developing new tools and approaches for DH analyses on numismatic collection data mainly based on the ontology framework of Nomisma.org. Another contribution of the DigiNUMA project is to create a generic semantic portal model, application, and LOD service for archaeological coin finds based on the "Sampo-model" [3] and Sampo-UI framework [4], a new part of the "Sampo" series of portals[6].

2 Data and Ontologies

As a case study the project concentrates on Viking Age (800–1150 AD) coins, selected owing to high degree of geographic circulation (therefore diversity in different national collections) of coins from this period in north-western and northern Europe and beyond from western Asia. DigiNUMA will mainly draw upon the existing numismatic data maintained by the Finnish National Museum Coin Cabinet and the Finnish Heritage Agency. In order to provide an international comparison, and to identify possible biases inherent in national numismatic datasets, DigiNUMA will also investigate English early medieval coin data from the British Museum, and data from the Corpus of Early Medieval Coins at the Fitzwilliam Museum in Cambridge, UK.

Digitisation of archaeological CH has advanced in many European countries in the recent years. Issues related to data harmonisation of archaeological data (alignment of dataset structures and object classifications) remain, however, a significant challenge for collating, studying and disseminating CH data at a transnational scale, as existing collections management practices and typologies typically descend from a broad variety of long-standing national or institutional practices and traditions. In this context numismatic data makes an excellent case study in LOD data harmonisation owing to a strong existing foundation of internationally shared typological practices. Nomisma.org was started by the American Numismatic Society in 2010, to facilitate the presentation of numismatic concepts using LOD [9] [2]. More than 30 different institutions have provided datasets[7] for the project.

Nomisma.org includes ontologies for many different aspects of numismatic data such as mints and rulers, and even deities depicted on the coins. Currently Nomisma.org mostly includes data related to classical era owing to the breadth of international scholarship on Greek and Roman numismatics, but expanding the ontologies to cover medieval era is planned. DigiNUMA project will aim to be a part of this process, and share the ontologies created with Nomisma.org.

[4] https://ariadne-infrastructure.eu
[5] http://Nomisma.org
[6] Sampo portals: https://seco.cs.aalto.fi/applications/sampo
[7] See: http://Nomisma.org/datasets

The ontology work done within the DigiNUMA project will be limited mainly to concepts relevant to the available Finnish data.

3 Using the LOD Service and Applications

As far as is possible, the research data will also be opened as a data service, but some elements of the data may not be possible to make public due to rights issues. Harmonized data together with ontologies, published as a LOD service on the Linked Data Finland platform,[8] is used in three ways: 1) the data can be filtered, uploaded and reused in external DH tools and applications. 2) The SPARQL endpoint can be used for data analyses using tools such as the YASGUI editor[9] and Jupyter Notebooks. 3) A new semantic portal "CoinSampo" is being developed on top of the SPARQL endpoint that can be used without programming skills. It demonstrates how the LOD service can be used effectively in application development. All data and software developed in DigiNUMA will be published openly using the CC BY 4.0 license, whenever this is allowed by the original data providers' copyright.

The CoinSampo web application is currently in early development, and is based on the Sampo-UI framework[10]. While the application is mainly developed for the Finnish data, using an international Nomisma.org ontology makes it possible to easily use the application to work with other data as well.

To test the early prototype of CoinSampo application, and it's applicability to international data, we have used the Seleucid Coins Online[11] dataset created by the American Numismatic Society. This dataset, like many others, can be downloaded from the Nomisma.org website and depicts coin types of the ancient Seleucid Empire (312 BC to 63 BC). As an example of the application prototype, Fig. 1 shows the relative numbers of coin denominations associated with the mint of Susa as a pie chart.[12] It is easy to see that most coin are of denomination "Tetradrachm". This can be very quickly compared to numbers from other mints.

Importantly, *DigiNUMA* seeks to develop applications for types of data analysis and visualisation that have been previously largely inaccessible without training in programs such as R or Python. Three examples are given. First, medieval numismatic data is typically "fuzzy data" in terms of its dating, with coins being dated within a range of possible dates of manufacture. In this context it is difficult to construct precise chronological charts displaying variation in coin observations across time using common statistical tables. Aoristic analysis is a temporal analysis technique for creating an aggregate picture of a set of temporally imprecise data [7]. The planned application assigns the fractional probability of each observation in a dataset falling into a given temporal bin,

[8] https://ldf.fi
[9] http://yasgui.triply.cc/
[10] https://github.com/SemanticComputing/sampo-ui
[11] http://numismatics.org/sco
[12] The application uses the ApexCharts library https://apexcharts.com to create chart visualizations.

adding them together and depicting the results as a bar chart that shows probabilistic patterns of increase and decrease in observations across time. Second, the application will also include various maps useful for analysis. Coin findspot data yields critical information on historical economic patters, yet, where findspots are numerous or clustered, simple point-based are inefficient for their close study. A kernel density estimation based "heat map" yields more useful map views [1]. Third, historic coin circulation can be visualised with radial diagrams that connect mint locations and coin findspots. On technical level, we use existing open source libraries[13] and the existing faceted search framework and modules of Sampo-UI to implement various kinds of visualizations. A researcher or hobbyist can use CoinSampo's faceted search functionality to create these kind of visualizations quickly and without technical know-how.

In an example use case a user interested in eleventh-century English coin economy could produce a chronological overview of all coins issued at a particular mint town, identify main periods of high output, and follow up by comparing spatial distributions to study changes in coin circulation patterns across time - possibly producing new information on growth and development of medieval regional economies. Should harmonised data be possible to add from different countries, such analysis could be extended to cover aspects of international trade.

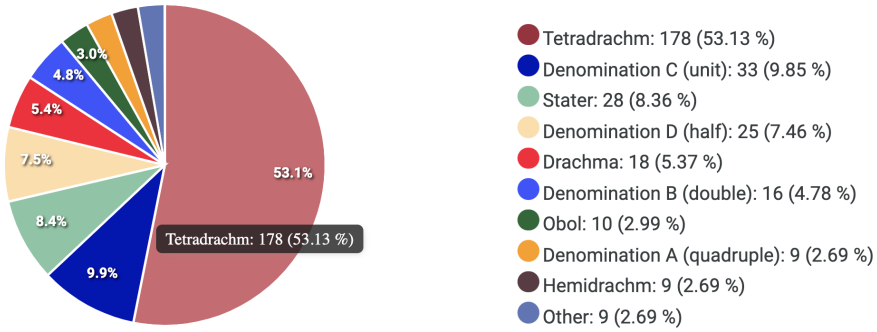

Fig. 1. An example of a visualization created with CoinSampo application showing coin denominations of Seluidic coins from the mint of Susa.

4 Discussion

In our experience, the current data in Nomisma.org datasets can sometimes have errors and inconsistencies, perhaps due to different data providers applying the

[13] For example ApexCharts: https://apexcharts.com

model in slightly differing ways. The ontology infrastructure is moreover currently limited for the post-classical era. However, the fundamental fact that a series of international datasets, arriving from institutions with divergent collections managements histories and policies, have nevertheless been described using a similar ontological framework makes it possible to reuse applications created in the DigiNUMA project for analysis and dissemination of this diverse body of numismatic data, with only slight alterations. The shared ontology infrastructure also makes creating new data easier. This demonstrates the significant potential of LOD and data harmonisation approaches in bringing together and creating added value from international and traditionally heterogeneous CH material.

A significant quantity of archaeological CH data in many European countries is today generated by the public. This is but one aspect that plays into the need to increase the accessibility of CH data, including its interoperability in scientific analysis internationally but also in lowering the threshold for anyone to discover, learn and create new knowledge, as a critical issue in European heritage management. To this end, *DigiNUMA* will lay the groundwork for future collaboration, secure international and interdisciplinary networks between Finnish and European partners, and develop solutions to current challenges.

Acknowledgements The project has been funded by the Jenny and Antti Wihuri Foundation. CSC – IT Center for Science has provided computational resources.

References

1. Bevan, A.: Spatial methods for analysing large-scale artefact inventories. Antiquity **86**, 492–506 (2012). https://doi.org/10.1017/S0003598X0006289X
2. Gruber, E., Meadows, A.: Numismatics and linked open data. ISAW Papers **20.6** (2021), http://hdl.handle.net/2333.1/q83bkdqf
3. Hyvönen, E.: Digital humanities on the semantic web: Sampo model and portal series. Semantic Web – Interoperability, Usability, Applicability (2022), submitted
4. Ikkala, E., Hyvönen, E., Rantala, H., Koho, M.: Sampo-UI: A full stack JavaScript framework for developing semantic portal user interfaces. Semantic Web – Interoperability, Usability, Applicability **13**(1), 69–84 (January 2022). https://doi.org/10.3233/SW-210428
5. Leahy, K., Lewis, M.: Finds Identified: Portable Antiquities Scheme. Greenlight Publishing, Witham (2018)
6. Oksanen, E., Rantala, H., Tuominen, J., Lewis, M., Wigg-Wolf, D., Ehrnsten, F., Hyvönen, E.: Digital humanities solutions for pan-european numismatic and archaeological heritage based on linked open data (2022), submitted
7. Orton, D., Morris, J., Pipe, A.: Catch per unit research effort: Sampling intensity, chronological uncertainty, and the onset of marine fish consumption in historic london. Open Quaternary **3**, 1–20 (2017). https://doi.org/10.5334/oq.29
8. Richards, J., Niccolucci, F. (eds.): The Ariadne Impact. Archaeolingua, Budapest (2019). https://doi.org/10.5281/zenodo.3476712
9. Wigg-Wolf, D., Duyrat, F.: La révolution des linked open data en numismatique : Les exemples de nomisma.org et online greek coinage. Archéologies numériques **17** (10 2017). https://doi.org/10.21494/ISTE.OP.2017.0171