This PDF is a simplified version of the original article published in Internet Archaeology. Enlarged images which support this publication can be found in the original version online. All links also go to the online version.

Please cite this as: Wigg-Wolf, D., Pfeiffer, A.-L. and Tolle, K. 2023 Aggregating Coin Find Data to the ARIADNE Portal. Challenges of using the Getty AAT for a specialist domain, Internet Archaeology 64. https://doi.org/10.11141/ia.64.6

Aggregating Coin Find Data to the ARIADNE Portal. Challenges of using the Getty AAT for a specialist domain

David Wigg-Wolf, Anna-Lisa Pfeiffer and Karsten Tolle

Integration of data on coin finds into the ARIADNE portal presents a number of challenges, partly related to the use of the Getty AAT as a controlled vocabulary, in particular its incomplete coverage and hierarchical structure. Further problems arise from the fact that the majority of the data are being provided not from primarily numismatic projects and institutions, but rather from disparate archaeological resources that integrate a wide range of artefacts and records. As a result, they often focus less on the peculiarities of using the AAT for coin finds and the difficulties that arise from its use. This article illustrates how this can lead to quite disparate and inconsistent mappings of data to the ARIADNE portal.

Potential solutions such as aligning the AAT to the established vocabulary of Nomisma.org, or even implementing the latter in the portal, as well as implementing a standard mapping for coin finds, are discussed. Also addressed are the possibilities for complex, granular searches using the SPARQL endpoint in the ARIADNElab Virtual Research Area.

1. Introduction

As mass-produced, more or less standardised, objects coins are ideally suited to digital applications, whether statistical analysis, recording in databases or in linked open data (LOD) applications. The Römisch-Germanische Kommission (RGK) des Deutschen Archäologischen Instituts in Frankfurt am Main, Germany, was involved in ARIADNEplus as a data provider, contributing data on coin finds from the database Antike Fundmünzen in Europa (AFE), and in particular the installation of

the database at the RGK, AFE-RGK. As such, AFE is something of an exception within the ARIADNE consortium, being dedicated solely to coins, in contrast to other data providers who, if they contribute numismatic data, do so as only one element of a broader spectrum of archaeological information. Coin finds are generally not a central focus of their work, even if they can be numerous. For example, of the 1,069,896 records provided by the Portable Antiquities Scheme of England and Wales, 503,276 are publicly accessible records of coins.

As numismatists specialising in coin finds, at the RGK we had expectations and demands on the process and form of data modelling for the ARIADNE portal, as well as on the search possibilities, that other providers and non-specialist users will not normally have. We were looking at finding ways of implementing the more granular searches for coins that numismatists often want, but which were not the primary intention of the portal

2. Antike Fundmünzen in Europa (AFE) as a Data Provider in ARIADNE

The AFE database that contributed data can look back at nearly 40 years of development, from early days in the late 1980s as a dBase III+ application, NUMIDAT, for the Fundmünzen der Antike (FdA) project at the Akademie der Wissenschaften und der Literatur, Mainz (Figure 1). From the outset the database was designed to record coin finds with as simple as possible a schema, but one that suited the Fundmünzen der römischen Zeit in Deutschland (FMRD) series of publications of finds of ancient coins from Germany. Subsequently the database was modified and extended in versions that ran with FileMaker pro and later Microsoft Access.

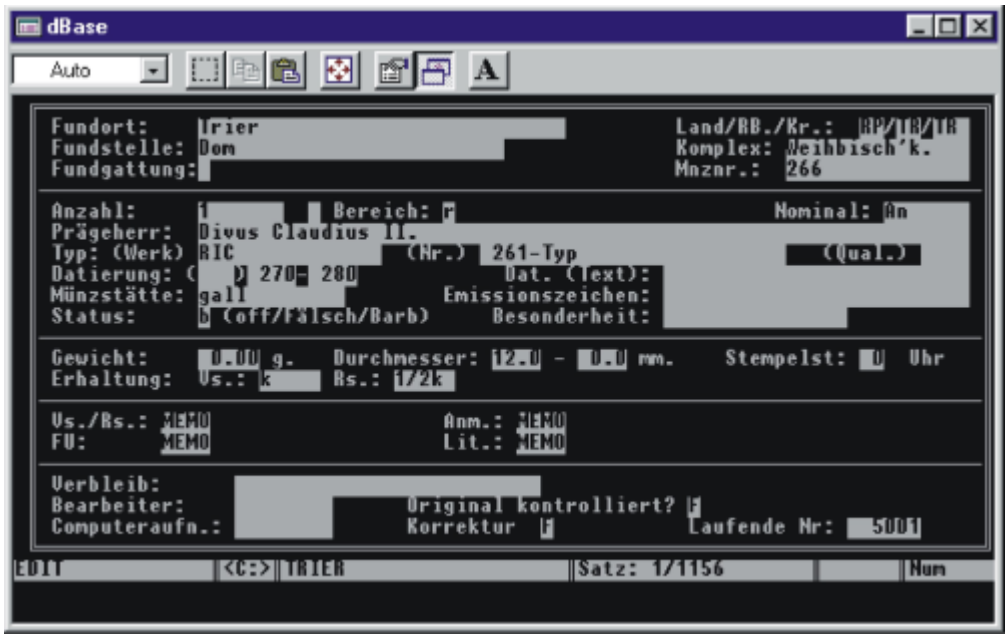

Figure 1: Screenshot of NUMIDAT, a dBase III+ database for recording coin finds

A major change came in 2009, when cooperation began between the RGK and Karsten Tolle from the Database and Information Systems department (DBIS, now the Big Data Lab) of the Goethe University Frankfurt to create an open source, web-based version of NUMIDAT. The database was renamed Antike Fundmünzen in Europa (AFE) and from the outset was integrated into the numismatic LOD world of the internationally recognised project Nomisma.org, as well as the various components of the digital world of the Deutsches Archäologisches Institut, the iDAI.world (Figure 2). AFE is a standard relational database, but employs a controlled vocabulary with drop-down thesauri for individual fields. The entries in the tables for the thesauri are linked where possible with concepts in Nomisma.org (Figure 3). This facilitates the creation of RDF that uses the Nomisma.org controlled vocabulary and ontology for uploading to the Nomisma.org SPARQL endpoint and portals such as Online Coins of the Roman Empire (OCRE). Similarly, bibliographic entries are linked to the iDAI.bibliography zenon.

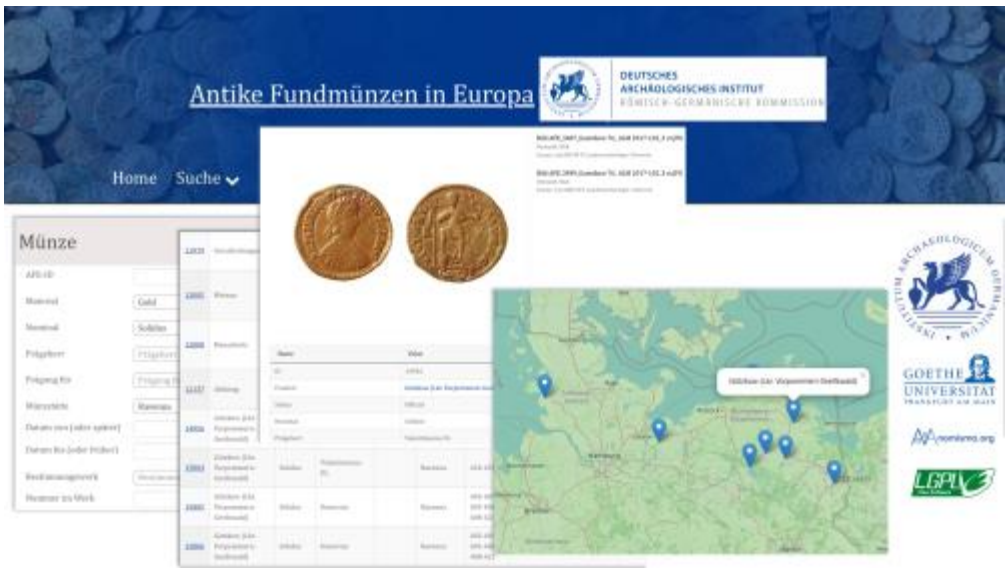

Figure 2: The public online frontend of AFE-RGK (http://afe.dainst.org/) (D. Wigg-Wolf, CC BY SA)

Id	Name	Name (en)	Nomisma	Active From	Active To	Ocre Short
33	Augustus	Augustus	augustus	-31	14	aug
48	Tiberius	Tiberius	tiberius	14	37	tib
52	Caligula	Caligula	gaius	37	41	gai
53	Claudius I.	Claudius I	claudius	41	54	cl
55	Nero	Nero	nero	54	68	ner
56	iulisch-claudisch	Julio-Claudian				
58	Galba	Galba	galba	68	69	gal
59	Vitellius	Vitellius	vitellius	69	69	vit
60	Vespasianus	Vespasianus	vespasian	69	79	ves
63	Titus	Titus	titus	79	81	tit
70	Domitianus	Domitianus	domitian	81	96	dom
73	Nerva	Nerva	nerva	96	98	ner
74	Traianus	Traianus	trajan	98	117	
77	Hadrianus	Hadrianus	hadrian	117	138	hdn
83	Antoninus Pius	Antoninus Pius	antoninus_pius	138	161	ant
88	Marcus Aurelius	Marcus Aurelius	marcus_aurelius	139	180	m_aur
93	Lucius Verus	Lucius Verus	lucius_verus	161	169	m_aur
94	Commodus	Commodus	commodus	175	192	com
97	Septimius Severus	Septimius Severus	septimius_severus	193	211	ss
100	Caracalla	Caracalla	caracalla	198	217	crl
102	Elagabalus	Elagabalus	elagabalus	218	222	el
103	Severus Alexander	Severus Alexander	severus_alexander	222	235	sa
105	Maximinus I.	Maximinus I	maximinus_thrax	235	238	max_i
106	Gordianus III.	Gordianus III	gordian_iii	238	244	gor_iii
107	Philippus I.	Philippus I	philip_the_arab	244	249	ph_i

Figure 3: Excerpt from the table of issuers in AFE. The columns include the ID in Nomisma.org, as well as the abbreviation for the emperor used to create the IDs for entries in Online Coins of the Roman Empire (OCRE)

The use of a controlled vocabulary is one of the advantages of AFE, allowing the full incorporation of the dataset into the numismatic LOD world. But that this can be done is partly because AFE is a purely numismatic database and the developers are themselves members of the Scientific Committee of Nomisma.org. Repositories and projects covering a wider range of archaeological materials and documentation will often not have the time and resources to implement controlled vocabularies for all categories of materials. And not all archaeological materials are suited to granular controlled vocabularies such as that developed by Nomisma.org.

3. A Controlled Numismatic Vocabulary, Nomisma.org

Nomisma.org was initiated by Andrew Meadows and Sebastian Heath of the American Numismatic Society in 2010 to provide a 'common currency for digital numismatics' and a namespace for numismatic concepts. Initial work focused on coinages of the Greek world, but since then Nomisma.org has expanded to cover a wide range of fields of numismatics and has working groups for Greek, Roman, Roman Provincial, Iron Age and Medieval/Modern numismatics, as well as hoards and iconography. It provides stable digital representations of numismatic concepts according to the principles of LOD in the form of HTTP URIs that also provide access to reusable information about the concepts, along with links to other resources. It also maintains a formalised RDF Ontology and a data model for encoding concepts,

coins, typologies, hoards and other types of numismatic objects as LOD (Tolle *et al*. 2018; Gruber and Meadows 2021). An extensive guide to using the vocabulary and ontology, the Nomisma Cookbook, is under development and has been published on the blog of Nomisma.org and the DARIAH Digital Numismatics Working Group (Figure 4).

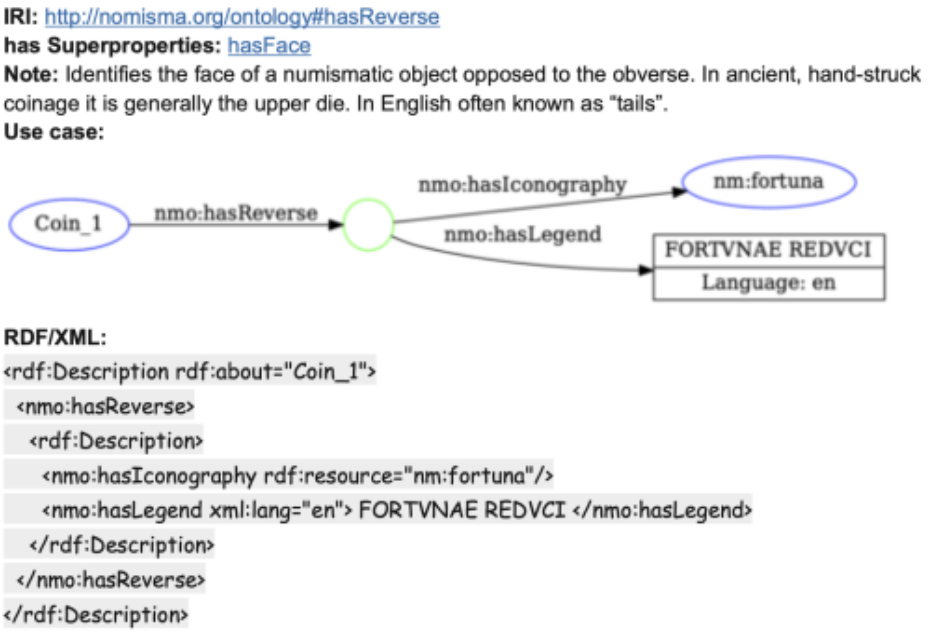

Figure 4: Information on the property hasReverse in the Nomisma.org ontology and an example of its use to model data (https://nomisma.hypotheses.org/1750)

The vocabulary and ontology are designed to reflect the way in which numismatists work and think, rather than to impose an overarching or external model on the material. Indeed, the project has the philosophy that if there is a concept that is widely used in the numismatic community, then it should be accommodated. This simple approach has resulted in an easily usable set of concepts that are now widely employed in numismatics and beyond.

The *ARIADNE WP15 Study: Towards a Web of Archaeological Linked Open Data. Version 1.0* (2016) (Debole *et al*. n.d.) cites Nomisma.org as an example of good practices from which the LOD community could benefit. A number of online typological portals, or virtual union catalogues, based on the vocabulary and ontology have been published that cover various aspects of Greek, Roman and medieval numismatics (for a list of such portals see https://fundmuenzen.org/links/). Other international projects such as Coin Hoards of the Roman Empire (CHRE) at the Ashmolean Museum, Oxford, are integrated into the numismatic LOD world by employing the concepts of Nomisma.org.

4. Mapping to the AAT

It was against this background that the RGK began to model coin find data for integration into the ARIADNE portal. A particular aspect of this task was the mapping of numismatic data to the controlled vocabulary of the Getty Art and Architecture

Thesaurus (AAT) that is employed for the portal. However, it soon became apparent that modelling data on coins with the Getty AAT is problematic for a variety of reasons. First and foremost, the AAT only covers a limited range of fields of numismatics (AAT: *coins by origin*), with a strong focus on the coinage of the ancient Mediterranean and medieval/modern Europe, with only sparse coverage of other fields such as Asia. The concepts listed consist mainly of names of denominations and a range of numismatic peculiarities (AAT: *coins by form or technique*) such as *overstrikes* or *punch-mark coins*. As a result, the AAT is not universally applicable. Byzantine coinage, for example, is barely covered: of the Byzantine denominations only *solidi* and their fractions (*semisses* and *tremisses*), *siliqae* and *miliarensia* are included, and then probably because they were also important elements of the late-Roman system. The bronze coinage and the entire later Byzantine system are absent.

But significant problems for the actual use of the AAT in the portal are also caused by the hierarchical nature of the thesaurus (Figure 5), in particular the way it was initially (not) implemented in the portal and how it has been used for mapping by individual data providers. Until the portal was modified to include the hierarchy at the end of 2022, a search on an AAT term only returned coins that were mapped directly to the term, but not to its children in the hierarchy. Thus, since most providers of numismatic content only map their coins to one single AAT concept, rather than to several – e.g. both parent and child – coins that were mapped to *denarii* were not displayed on a search for Getty AAT Subjects: *Early Western World coins*. This problem has now been solved by implementing the hierarchy of the AAT in the portal, so that all children of *Early Western World coins* are displayed, including *denarii*. Nevertheless, since most providers only map to one concept, generally preferring the parent term to the child, the AAT is not useful for digging down below the more general level of *coins (money)*. Searches on the children of an AAT term will not provide statistically relevant figures for the number of coins corresponding to each child actually present in the portal, since the children are not uniformly mapped to the AAT by all providers.

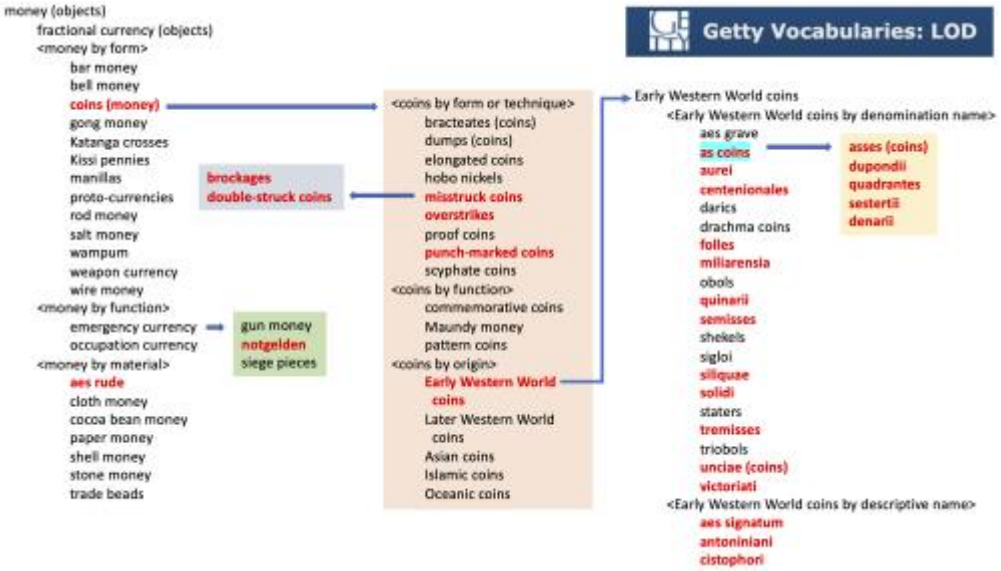

Figure 5: The hierarchy of the AAT thesaurus for concepts pertaining to Roman and Greek coinage

To dig down deeper it is necessary to use the Original subject search, that is the concepts used by the providers within their systems. As will be shown below, however, Original subject does not provide a controlled vocabulary. There is no consistency across providers; indeed providers are not always consistent internally.

Even now that the hierarchy has been implemented in the portal, and children of AAT concepts are also returned in search results, a further, potential problem is caused by some related concepts being siblings, rather than parent or child. In particular, the division of c*oins (money)* into *coins by form or technique* and *coins by origin* means that a range of numismatic phenomena frequently found on ancient coins such as *overstrikes* or *punch-marked coins* will not be found in a search for *Early Western World coins*, the former being a child of *coins by form or technique*, the latter of *coins by origin*. A search using Getty AAT subjects: *denarii* filtered to Getty AAT subjects: *overstrikes* would require the data entry being mapped to both *overstrikes* and *denarii*. But as noted above, numismatic data providers normally only map to one AAT term, if they map to the AAT at all.

More problematic is the assignment of *emergency currency* (AAT: 'Note: Money issued under abnormal financial conditions, such as wartime or economic panic') as a cousin of *coins* (*money)*. The former could be used for the numerous imitations of Roman coins, known in British numismatics as (barbarous) radiates, which were produced in the late 3rd century CE during a period of inflation and a chronic shortage of official coins. Both terms would only come up together on a search for *money (objects)*. The same applies to *aes rude*, an early form of bronze money in Central Italy, which immediately preceded the earliest Roman coinage and which is a child of *money by material*. Probably neither of the concepts *emergency money* and *aes rude* are actually used at present by any content providers, but since at the time of writing there were problems in searching for certain AAT terms in the portal, it was not possible to check if this is indeed the case. Certainly, *aes rude* is not included among the Original subjects in the portal.

A very particular problem arising from the hierarchical implementation of the AAT in the portal is caused by a case where the confusing labelling in the AAT has led to incorrect use of a term. The coinage of the Republican and Augustan coinage systems is hierarchically arranged in the thesaurus, with the individual denominations as children of the parent *as coins* (with the exceptions of *aurei*, *quinarii* and *semisses*, which are siblings of *as coins* in the hierarchy of the AAT. In the case of *semisses*, this is because of its dual use for late-Roman and early-Byzantine gold – see below – but why *aurei* and *quinarii* are siblings and not children is unclear). *As coins* is a term not usually employed in numismatics to describe this coinage as a whole, but it does reflect the fact that the Republican and early Imperial systems were based on the *as* as the basic denominational unit. The problem is that AAT defines *as coins* as 'Early Roman copper coins', even though the *as* and *quadrans* were the only copper coins in the system, the others being of brass, silver and gold. As a result, one provider has misunderstood the term and modelled all *asses* in their dataset to *as coins* rather than *asses (coins)*. Previously, before the implementation of the AAT hierarchy, this was not a problem as a search for *as coins* produced just the coins modelled as *as coins*. But now the search produces not only coins modelled directly as *as coins*, but also all children of the term. The 45,193 results of a search on *as coins* include not only *asses*, but also

19,381 *denarii*, 11,775 *sestertii*, 7787 *dupondii* and 21 *quadrantes*, all of which can then be displayed separately (Figure 6). However, although the result list also indicates that there are 6231 coins modelled as *as coins*, they cannot be displayed separately in the portal. There is no possibility of displaying only the coins mapped to the parent without those mapped to the children.

Figure 6: The list of AAT terms produced by a search 'Getty AAT subjects: as coins'

Here again, the only way to circumvent this is to dig down deeper using Original subjects. But as noted above, this brings its own problems as the individual data providers use different terminologies that are generally not compatible with each other and detailed searches across different providers are therefore not always possible. This can even be the case within the data of one provider. For example, the British Museum has what is properly a subset of the radiate coinage of the second half of the 3rd century AD: the Q radiates produced during the reign of the breakaway British emperor Allectus, which have their own Original subject, Q radiates. However, they will not be included in a search for Original subject: Radiate (antoninianus) since there is no hierarchical relationship between the two. The matter is further complicated by the fact that the RGK uses just 'Antoninianus', 'radiate' being a term used in English-language and not German numismatics. The RGK's coins are also mapped to AAT: *antoniniani*, which the British Museum does not use.

Beyond the problems arising from the AAT hierarchy, its implementation and use, a number of rather different problems are caused by the limited range of concepts and fields of numismatics covered by the AAT. There is also a degree of confusion as a result of the AAT using in at least one case the same term for different

denominations. Thus *semisses* covers both the gold half *solidi* of Late Antiquity and the early Byzantine period, as well as the half bronze *asses* of the Roman Republic and early Empire. The use of the plural form by the AAT for denominations is unfortunate here, as the two terms *can* have different singular forms: *semis* rather than *semissis* is more usually used for the half *as*, so that a differentiation would have been possible. As a result, a search for Getty AAT subjects: '*semisses'* in the ARIADNE portal returns 55 hits, of which two are late-Roman gold coins, the rest Republican/Imperial bronzes. While the expert human user will immediately recognise this and use the relevant instances for their particular case, uncritical machine reuse that does not process the relevant metadata could lead to 22.5 times too many late Roman/Byzantine gold *semisses* being recorded in reuse.

Coverage of the coinage of Late Antiquity by the AAT is generally problematic, not least because our understanding of the coinage of the later Roman Empire is patchy. In contrast to the Republic and early Empire, for which we have good knowledge, for Late Antiquity the ancient sources give us names for some denominations, in particular for bronze coins, but we do not know which coins they actually refer to. The AAT includes only two of these names for denominations, both of which are used in German academic tradition, *folles* and *centenionales:* the first for bronze struck until 348, the second for bronze struck subsequently. This is problematic because the terms are not employed in English-speaking numismatics, where a different terminology is used. Only the RGK, as an institution in the German tradition, uses the AAT terms. Furthermore, quite apart from the question as to whether the attributions used by the AAT are correct, a number of other important terms in the German tradition are not included
– *Halbfollis*, *Maiorina*, *Doppelmaiorina* and *Halbcentenionalis* – so that only some of the late-Roman bronze coins can be mapped to the AAT. Coverage for the late Roman precious metal is better, with the main
denominations *solidi*, *siliquae* and *miliarensia* included, but the divisions of the latter two are again missing, as are a number of rarer fractional gold units.

The 256 bronze coins dating to the period 293-348 are accordingly mapped as *folles* by the RGK, while the 101,621 equivalent coins published in the portal by the British Museum are modelled as Original subject: nummus (ae1 – ae4). However, since it is not possible to create a search in the portal that combines the results of the AAT term used by the RGK and the Original Subject term of the British Museum, and since the RGK coins are mapped as Original subject: follis and not Original subject: nummus (ae1 – ae4), it is not possible to carry out a targeted search for the bronze coins of 293-348 that includes the coins of both publishers in the result. This is only possible with a more general search for 'coin' filtered to the date range 294–348, but which will also include gold and silver, as well as a range of other coins that are more generally dated to a broader period (e.g. AD 43–410 in the case of a number of coins from the British Museum). However, as is discussed below, the SPARQL endpoint in the ARIADNElab Virtual Research Area does facilitate such granular, complex searches.

However, this search also demonstrates the potential of the use of the AAT in ARIADNE for quality control: it reveals a number of medieval coins that have been incorrectly dated; for example a half *groat* that had been erroneously dated (due to copy and paste?) to 348–350

(https://finds.org.uk/database/artefacts/record/id/961799), or another groat dated to 182–1583 as the result of a typing error (https://finds.org.uk/database/artefacts/record/id/47019).

The AAT's definition of the term *folles* also provides a potential source of confusion. It is defined by Getty as 'Ancient Roman copper coins of the 3rd and 4th centuries, replaced by centenionales'. While, as noted above, the term is indeed used in German academic tradition in this way, it is more generally used for the large Byzantine bronze coins with a value of 40 *nummi* introduced by the emperor Anastasius in 498. The definition by Nomisma.org, which uses the singular 'follis', notes the potential confusion: 'Name for Byzantine bronze coins worth 40 nummi. *Sometimes used for the Roman billon coins of the Diocletianic reform and of the first half of the 4th century AD, although there is no concrete evidence for the attribution*. The Byzantine coins bear the letter M resp. numeral XXXX as a mark of value' (italics by the authors). Fortunately, the only data provider using the term follis for the Byzantine coins, the British Museum, refers to them as 'copper follis (40 nummi)', thus avoiding any confusion.

Much like the problems arising from the hierarchical position and use of *as coins* discussed above, the AAT terms cannot be used for any purposeful granular search of coins, since only a small proportion of them can be mapped to the AAT. Nor is the AAT terminology widely used for providing detailed information on coins in the ARIADNE portal. Instead, the various contributors resort to a variety of terminologies of their own to describe coins in more detail, and these are mapped to 'Original subject'. The problem here, however, is that without a standard vocabulary the individual terminologies are generally not compatible with each other and detailed searches across different providers are not always possible. As was shown above, this can even be the case within the data of one provider. But these problems are not so much a fault of the ARIADNE portal, for the portal itself is primarily designed for more general queries. What is more, as noted above, the SPARQL endpoint in the ARIADNElab VRE does allow for more complex, granular searches, although this requires good knowledge of SPARQL, as well as of the various vocabularies employed by the different data providers.

However, problems can arise when detailed data and mapping are available and implemented in the portal, especially if searches are carried out uncritically. A final example can serve to highlight the problem. A search for 'Search: byzantine; Original subject: coin' produces 130 hits, of which 9 are mapped to AAT-*solidi*, 7 from the dataset of the RGK, 2 from the British Museum. A search for 'Dating: byzantine; Original subject: coin' also produces 130 hits, which superficially appears consistent. However, the hits now include 19 coins mapped to AAT-*solidi*, 15 from the RGK, 4 from the British Museum. The discrepancy results from the fact that the first search looks for the word 'Byzantine' in the data submitted to ARIADNE, but not for 'Dating: byzantine', and if the word is not included then the item will not be returned as a hit. As a result, those coins including the description 'derived from Byzantine prototype' but are in fact Western medieval coins are included, while many actual Byzantine coins are excluded because their description does not include the keyword.

5. Lessons Learned and Future Possibilities

In order to implement granular searches in the portal, a standardised, controlled vocabulary would be necessary. The above examples demonstrate the difficulties of using the Getty AAT in this way within a specialist domain, difficulties that can be divided into two aspects. The first relates to the AAT itself and its hierarchical structure, which does not always relate terms to each other in a useful way (e.g. *emergency money* and *aes rude*), or else does not place them correctly within the hierarchy (e.g. *quinarii* and *aurei*). The implementation in the portal can also be problematic. The second aspect arises from the incomplete, sometimes inconsistent, mapping of data to the AAT in the portal. This is understandable, given that most data providers are not specialist numismatic projects, are often contributing data from a wide range of archaeological materials, and do not always have the capacity to map in detail. Granular searches in the portal can be carried out using Original subjects, but as this is not a uniform, controlled vocabulary, there is no consistency across data providers, sometimes not even within the data of one provider (e.g. one provider uses both 'half siliqua' and 'half-siliqua').

In an LOD world of automatic or machine reuse of datasets, this poses risks if data are reused uncritically, without an understanding of the limitations of the underlying mappings to the vocabulary, whether AAT or Original subjects.

Clearly the use of a more complete controlled vocabulary would greatly enhance the potential of data on coin finds within the ARIADNE portal, and as mass-produced, more or less standardised objects, coins are inherently suited to such an implementation. The Nomisma.org vocabulary could be a candidate, offering extensive coverage of the terminology of a wide range of fields of numismatics, and reflecting, as it does, the ways in which numismatists and archaeologists think about coins, as well as how numismatic data is widely structured in databases. Alternatively, the Getty AAT could perhaps be modified and extended to match the vocabulary of Nomisma.org, at least for the areas relevant to the ARIADNE portal. This would have the additional advantage of implementing a hierarchical structure that at present is not supported by Nomisma.org. It would also avoid potential problems arising from the use of two different controlled vocabularies.

However, it should be noted that the ARIADNE portal is not primarily intended for such granular searches in a specialist field, and it was not possible to implement the Nomisma.org vocabulary within the framework of the ARIADNEplus project. But the work of mapping the data of the AFE-RGK database to the portal, and in particular the comparison with the mappings used by other projects, has led to a number of important insights into the problems of the use of a general controlled vocabulary such as the AAT to a specialist domain, and can suggest future ways ahead. For example, the common use of a simple, easily applied standard mapping for coins that employs the vocabulary of Nomisma.org or a revised AAT would facilitate granular searches across datasets, greatly empower the portal, and increase its attractiveness for a variety of user communities.

Acknowledgements

This work was supported by the European Commission under the H2020 Programme, contract no. H2020-INFRAIA-2018-1-823914 (the ARIADNEplus project). Thanks are due to project partners generally and to the editors for helpful comments. The views and opinions expressed in this article are the sole responsibility of the authors.

Bibliography

Debole, F., Meghini, C., Geser, G. and Tudhope, D. n.d. *D15.2: Report on the ARIADNE Linked Data Cloud*. http://legacy.ariadne-infrastructure.eu/resources-2/deliverables/d15-2-report-on-the-ariadne-linked-data-cloud/

Gruber, E. and Meadows, A. 2021 'Numismatics and linked open data', *ISAW Papers* **20**(6). http://hdl.handle.net/2333.1/q83bkdqf

Tolle, K., Wigg-Wolf, D. and Gruber, E. 2018 'An ontology for a numismatic island with bridges to others' in M. Matsumoto and E. Uleberg (eds) *Oceans of Data. Proceedings of the 44th Conference on Computer Applications and Quantitative Methods in Archaeology*, Oxford: Archaeopress. 103-8.