



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: Vassallo, V., Theodoridou, M., Felicetti, A. and Avgousti, A. 2023 Tools and Ontologies for the Aggregation and Management of Cypriot Archaeological Datasets, Internet Archaeology 64. <https://doi.org/10.11141/ia.64.10>

Tools and Ontologies for the Aggregation and Management of Cypriot Archaeological Datasets

Valentina Vassallo, Maria Theodoridou, Achille Felicetti and Avgoustinos Avgousti

This article focuses on the aggregation of Cypriot archaeological datasets, digitally archived in local repositories, into the ARIADNE portal. It considers, in particular, the development of an application profile for inscriptions and presents the integration of two collections, consisting of ancient coins and inscriptions carved on stones. It highlights the tools and ontologies developed for the aggregation and management of these digital resources, as well as the related pipeline and activities. The issues encountered are also presented, plus the solutions adopted and the successful results in the data aggregation of these collections into the infrastructure. Currently, thanks to the pipeline, and the semantic tools developed and used in ARIADNE, a collection of Cypriot medieval coins and a corpus of Ancient Greek inscriptions are now more widely accessible to the archaeological community.

1. Introduction

In recent decades, there has been a boost in the creation of digital archives and digital collections, both from public and private initiatives (Hawkins [2022](#)). On the one hand, this has generated an increase in the openness of culture and its dissemination, and consequently a longer-term solution to conservation and preservation. On the other hand, this situation has also created a dispersal of material into different databases and online repositories. Indeed, while digital archives provided online data access, services or tools were still missing, including data discovery through connections among data and repositories, data sharing or the possibility of data reuse, solutions that in recent years are highlighted as necessary by the FAIR principles (Wilkinson *et al.* [2016](#)). In the last few years, research communities identified these problems and started to create digital 'places' where



content could be standardised, unified, and linked. In particular, research e-infrastructures assumed the role of becoming those digital 'places', and started the transformation into more collaborative platforms for the aggregation and interconnection of data. Moreover, the increase in the digitisation or production of digital data and their different nature pushed the implementation of digital and semantic tools for analysis, management and standardisation (Vassallo and Felicetti [2021](#)). The cultural heritage field, and specifically the archaeological one, has been at the forefront concerning the use of digital tools for data documentation, and even development, for analysis, management and publication. The archaeological community, and notably [ARIADNE](#), started to address these issues in the last decade (Meghini *et al.* [2017](#)). The archaeological e-infrastructure set up and further development during the entire EU-funded period provides a solution to enable data providers, even small countries and institutions, to be seen, and to provide access to their national and local archaeological digital resources. Moreover, the portal, through the development of tools and services, facilitates discovery, replying to users' queries, and, at the same time, enables the creation of new enquiries, allowing the further development of research.

To date, the ARIADNE e-infrastructure has aggregated more than 3.9 million archaeological resources of numerous cultural institutions from several European and non-European countries (including the USA and Japan). The solutions developed within the infrastructure are aimed at meeting the FAIR principles, proposing a Knowledge Base built upon linked open data and semantic instruments, guaranteeing standardisation from the descriptive, temporal, geographic and thematic point of view, and tools for the exploration and analysis of the aggregated data (Richards [2023](#)).

This article focuses on the provision of Cypriot archaeological datasets, digitally archived in local repositories, to the ARIADNE portal. Under one of the sub-domains aggregated by the e-infrastructure, namely the application profile for inscriptions, this contribution presents the aggregation and integration of two collections, consisting of ancient coins and stone inscriptions (Richards *et al.* [2022](#)). In particular, the article highlights the tools and ontologies developed and used in the e-infrastructure for the aggregation and management of the resources, and the related pipeline and activities. It also presents the issues encountered, the solutions adopted and the successful results in the data aggregation of the digital collections into the e-infrastructure, as well as future perspectives.

2. State of the Art in the Field

As previously mentioned, the European landscape of data repositories for cultural heritage has flourished in recent years, with scope to solve first data digitisation and aggregation, and then the lack of standardisation, interoperability and services. For instance, in the first decades of the 21st century, there were several European-funded projects dedicated to the aggregation of cultural heritage data, mainly in museums, libraries and archives, facilitating their digital access and integration, as well as their content provision to [Europeana](#), the European cultural portal, including the [ATHENA](#) and [EDLocal](#) projects. These initiatives focused on improving search, retrieval and reuse, multilingual terminology management, SKOS export and



publication tools, and experimented with enriched metadata for their reuse, via [ATHENA Plus](#) and [Linked Heritage](#). Generally, these projects and the data repositories created were mainly aggregating mixed cultural heritage data where archaeology, if present, was just one of the areas of aggregation. During the period from 2010 to 2020, although mainly focused on the general public rather than researchers, and providing access for several purposes (e.g. reuse for education, tourism, and virtual exhibitions), new European projects started to be committed to the aggregation of archaeological data, including [CARARE](#) and [3D-ICONS](#). Moreover, these two projects included not only 2D but they also centred attention on the integration of 3D archaeological data, following an increasing interest in 3D documentation and digitisation occurring in the field. Indeed, during that period, many European countries started to create national or regional databases dedicated to the archiving of 2D and 3D archaeological data (e.g. sites, monuments, artefacts) and their management. Nevertheless, an assessment shows a varied landscape, with some countries taking an early lead in promoting digital archiving for archaeology, others staying relatively behind, and many that, despite the creation of digital repositories for management purposes, rarely providing public access (Richards *et al.* [2021](#); Jaillant [2022](#); Vassallo *et al.* [2023](#)). The situation was already highlighted some years ago by Meghini *et al.* ([2017](#) 3-5) who showed that the years of initiatives led to an increase in integrated datasets under national portals, either aimed at research or at digital preservation and open access, but mainly concentrated in northern European countries (e.g. [ADS](#) in the UK, [SND](#) in Sweden, and [Arachne](#) in Germany). The lead of the northern countries and the commencement of the ARIADNE activities pushed other states to follow this example and develop their national or inter-institutions archaeological repositories (e.g. the [Archaeology Database](#) of the Hungarian National Museum).

In this context, the contribution of ARIADNE to these distributed and varied initiatives was to recognise the necessity of an archaeological digital infrastructure to unify the different resources, while supporting their maintenance at the national level, where the legal responsibility can better guarantee the maintenance and future preservation of national archaeological datasets. In this way, ARIADNE provided integration through semantic interoperability, allowing cross-searching of the collected resources at a European and international level (e.g. USA, Argentina, Israel, and Japan).

In ARIADNE this is achieved through ontologies and tools guaranteeing data harmonisation and standardisation, accessibility to the different resources, their interoperability and findability. The solutions are based on linked open data and semantic instruments that guarantee standardisation from different points of view, including the descriptive, thematic, chronological and geographic, and tools developed for the exploration and analysis of the data.

The exponential growth of digital data in all the fields, such as life sciences, environmental studies and humanities led the related communities to find efficient data management solutions, e.g. [LifeWatch](#), [PANGAEA](#), [CLARIN](#), and [DARIAH](#). All these infrastructures, as well as the distributed repositories around Europe and beyond, develop and use different data models, ontologies and vocabularies, specific to the kind of data they need to describe. In fields such as Life or Environmental Sciences the semantic resources for describing, integrating, and



normalising datasets are, for instance [CERIF](#), [Darwin Core](#), [OBOE](#) and [EnvThes](#). In the cultural heritage field, many semantic solutions and tools now exist that allow data integration and interoperability. Core and domain ontologies are increasingly becoming the preferred tools to structure complex archaeological data. CIDOC CRM and its extensions, in particular, have recently been used in many data integration projects and initiatives, such as [ItAnt](#), [EAGLE](#) (integrating it with [TEI/EpiDoc](#), the semantic standard for encoding scholarly and educational editions of ancient texts), and [Epigraphy.info](#) that continued the work of EAGLE on ontologies and vocabularies. Indeed, in archaeological fields such as numismatics and epigraphy, there are several research communities dedicated to the development or enhancement of vocabularies, translation tools and search engines (e.g. EAGLE, [FAIR Epigraphy](#) and [Nomisma.org](#)). EAGLE developed several instruments for the epigraphic field, including an ontology, [vocabularies](#), a [translation tool](#) and a [search engine](#). The ARIADNE e-infrastructure with its development and use of semantic solutions and tools for collecting and managing archaeological data is aligned with the most important initiatives in the domain and fits in the wider landscape aimed at addressing the FAIR principles.

3. The Cypriot Collections

The Science and Technology for Archaeology and Culture Research Centre of the Cyprus Institute (STARC) has contributed to the aggregation of inscriptions within the ARIADNE research infrastructure. The collections are archived in repositories derived from the collaboration of national and local institutions aimed at digitally documenting and preserving the archaeology of Cyprus. Specifically, the Cypriot resources aggregated in ARIADNE comprise a collection consisting of ancient coins and a corpus of ancient Greek inscriptions (Figure 1).



Figure 1: Cypriot archaeological materials aggregated in the ARIADNE portal

The features and character of these collections and their chronological and geographical coverage were the motivation for integrating them into the archaeological e-infrastructure. Indeed, an explicit aim of ARIADNEplus was to increase the temporal and geographical range of archaeological resources in the portal and to extend its thematic coverage. The possibility to discover, access and



share Cypriot archaeological artefacts through the portal helped to make them more visible and better known both within and across regional and national borders. This is especially important because in the past Cypriot archaeology has been subject to export and dispersal to other countries as a result of several events. For instance, the legislation in force during the 19th-20th centuries allowed the appropriation and export of archaeological materials outside the country and could not effectively regulate the trafficking of finds from illicit excavations (Stanley-Price 2001). Moreover, the Turkish occupation of 1974 had an impact on the country's archaeology, with illicit trafficking, looting and the consequent dispersion of Cypriot antiquities around the world. The integration of Cypriot archaeological datasets in the infrastructure facilitates their global sharing, interoperability with resources archived elsewhere, and their digital preservation.

3.1. The Cypriot medieval coin collection

The coin collection aggregated in the ARIADNE portal consists of [Cypriot medieval Coins](#) published in the Cyl DIOPTRA, a digital library dedicated to the rich Cypriot cultural heritage that covers topics ranging from archaeology to anthropology to art history and culture (Figure 2). The numismatic collection documents the use of currency on the island during the Middle Ages, specifically during the Cypriot Frankish (1192-1489) and Venetian periods (1489-1571), therefore spanning the 12th to the 16th century.

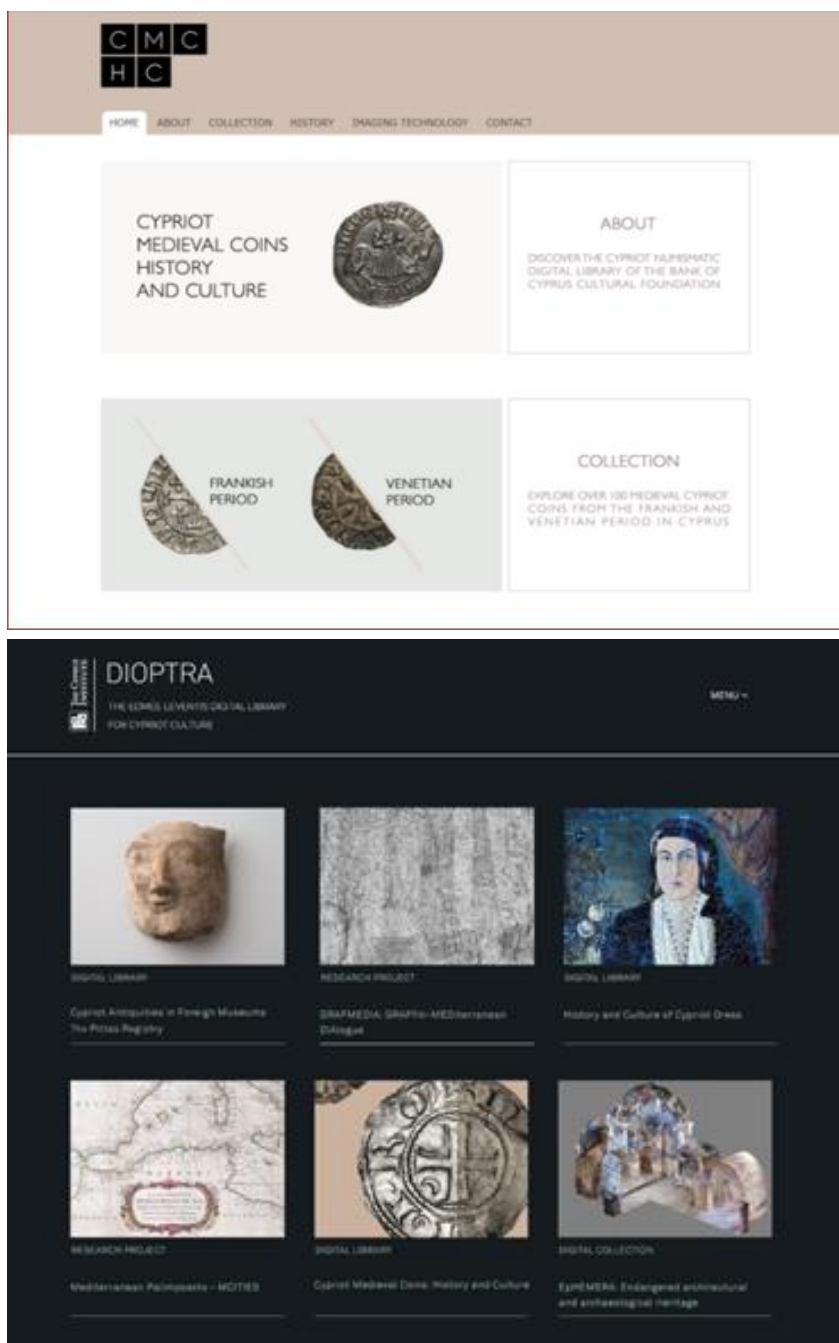


Figure 2: The Cypriot medieval coin collection in DIOPTRA

The creation of the numismatic digital collection was the result of a combined effort between the Cyprus Institute and the Bank of Cyprus Cultural Foundation, whose Museum of the History of Cypriot Coinage in Nicosia curates the physical coin collection. For the development of the web-based platform, an open-source content management framework, Drupal 7, was employed. The target was to create a platform specifically designed for digital numismatic collections: to achieve this goal, various advanced technologies and semantic tools for the management, analysis and description of these specific artefacts were utilised, such as advanced search, facets, image search, embedded interactive maps, metadata, taxonomies, XML export, and bibliographical references. Beyond providing access to the coinage of the period, the digital platform offers interactive exploration of the coins through

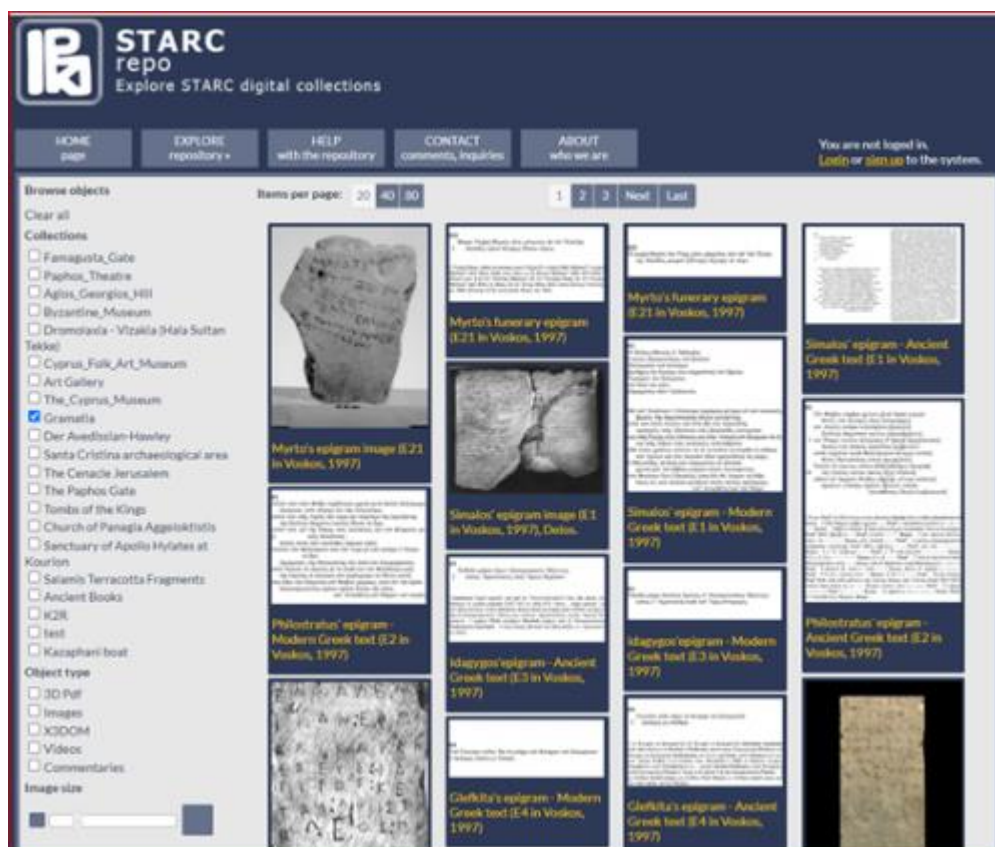


Figure 4: The Archaia Kypriaki Grammateia in the STARC Repo

The Archaia Kypriaki Grammateia consists of a corpus of Ancient Greek Cypriot inscriptions collected in physical volumes (Voskos *et al.* [1995](#)). The ancient texts corpus includes a wide range of literary genres (e.g. epic, lyric and dramatic poetry, epigrams, prose, medical and philosophical texts) inscribed on stone. The entire corpus covers Cypriot literary production for a time span of almost thirteen centuries (7th century BCE-6th century CE). In particular, the digital data archived in the STARC Repo and integrated into the ARIADNE portal consist of Cypriot epigrams inscribed on stones found in Cyprus and elsewhere (mainly in Greece) and conserved in several locations. The collection gives access to different resources such as the transcribed Ancient Greek text, the translated Modern Greek text, the scholia (commentaries), the image of the archaeological object (that is the inscription support), and, in some cases, its 3D representation. Every item of the dataset is described using an extension of the STARC metadata schema (Ronzino *et al.* [2012](#)), a CIDOC-compliant metadata schema, extended to thoroughly describe the various aspects of an inscription (Vassallo *et al.* [2013](#)). This collection also has a double value as it enriches another initiative for digital text curation dedicated to the epigraphy community and is available for education purposes (Pitzalis *et al.* [2012](#)) published in the [DIOPTRA](#) digital library (Figure 5).

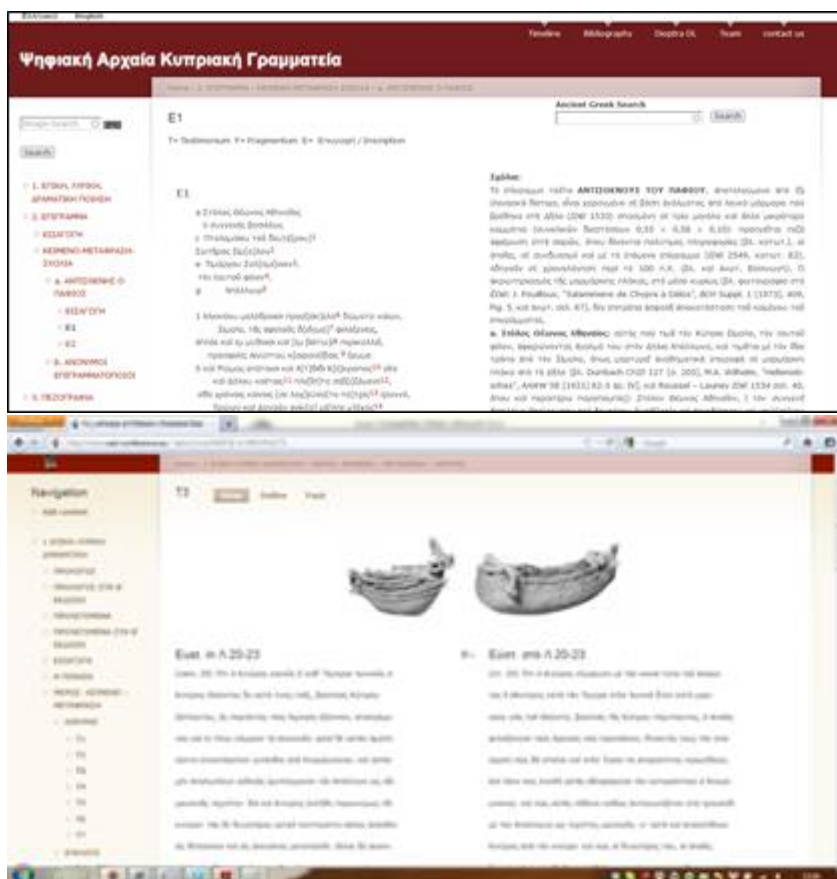


Figure 5: The Archaia Kypriaki Grammateia published in DIOPTRA

4. The Aggregation Pipeline: from Cypriot repositories to the ARIADNE portal

Within ARIADNEplus, in order for the digital collections to be aggregated and integrated into the portal, a pipeline consisting of specific steps and mandatory criteria in the data description had to be followed. This series of constraints in the aggregation procedure was introduced in order to guarantee data standardisation, ingestion, and to address the FAIR principles (Bardi *et al.* 2022; Hollander *et al.* 2018). Specifically, for the aggregation of the two Cypriot collections, the steps consisted of metadata extraction and cleaning, their enrichment, including the terminological, thematic and chronological standardisation, and the ontological harmonisation for the publication in the portal (Figure 6). The pipeline can be summarised as follows:

1. Metadata extraction, cleaning and improvement
2. Enrichment
 - a. Chronology standardisation
 - b. Terminology standardisation
3. Ontological harmonisation
4. Publication, semantic browsing and search



In order to carry out these steps, several semantic solutions and tools for the data preparation, aggregation, and management of the archaeological datasets - some specific to the inscribed artefacts – were integrated and applied:

- Metadata schemas
- PeriodO
- Vocabulary Matching Tool
- CRMtex ontology for inscribed artefacts
- X3ML Toolkit
- Staging portal for preliminary data control and evaluation
- Semantic search demonstrator

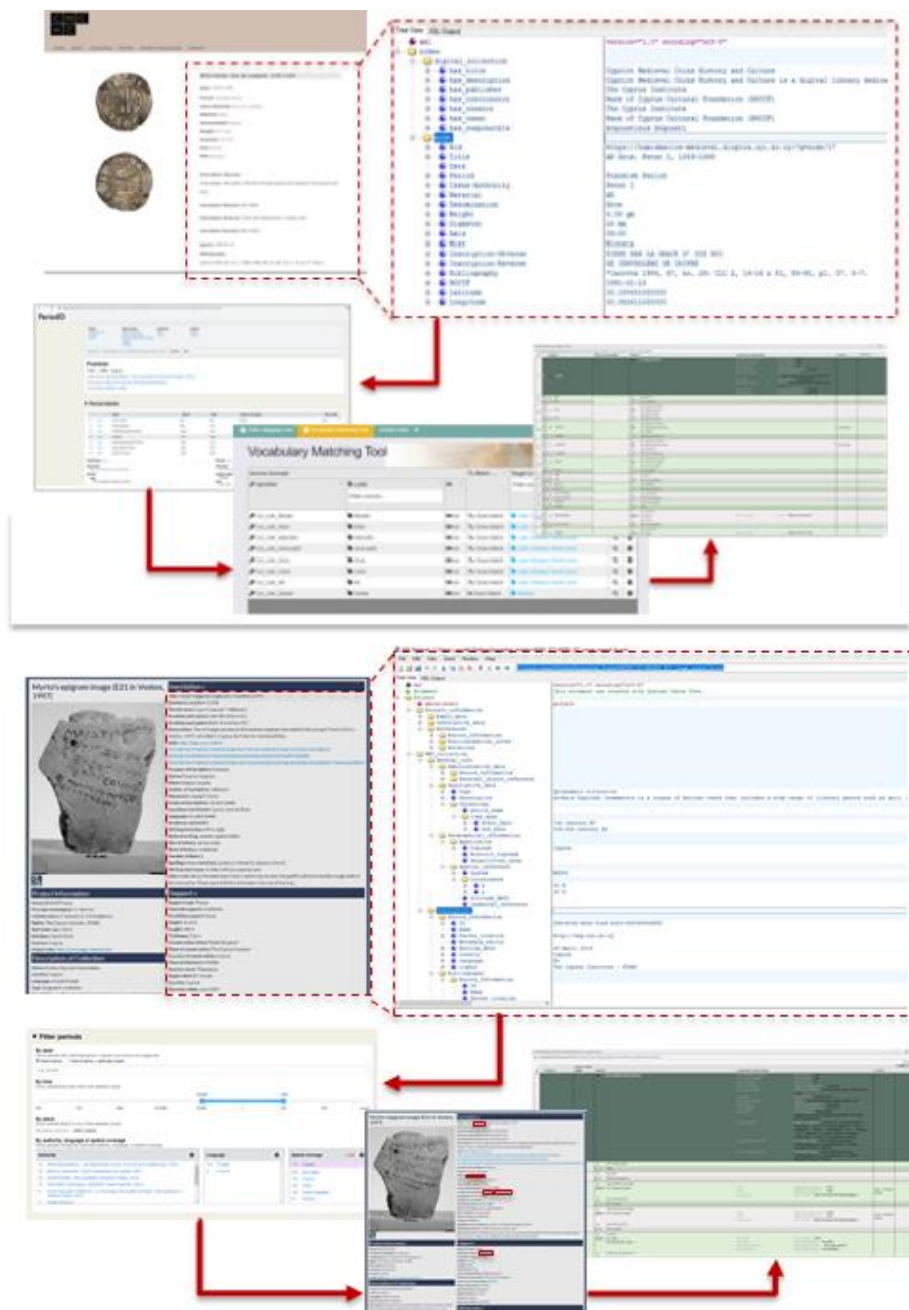


Figure 6: The schematisation of the aggregation pipeline for a) the coin collection and b) the inscriptions corpus



Metadata extraction, cleaning and improvement

The aggregation of the two Cypriot collections started with the extraction of the metadata schemas used to describe their items in the source repositories. The use of metadata schemas to describe the collections is one of the semantic tools necessary to guarantee a first step towards the standardisation of the related information. In the case of the coin collection, the metadata used is straightforward, including a few fields usually employed in the traditional description of the physical coins at the museum where they are conserved. On the one hand, this linear structure helps to describe such objects easily. On the other hand, many things and much information were missing for a proper alignment with the ARIADNE ontology, and for compliance with the mandatory or desirable fields required by the aggregation pipeline. Some cleaning was also necessary. The extracted metadata consisted of a single XML file; therefore, this needed to be subdivided into different nodes for each resource, as required by the ARIADNE aggregation rules. The analysis of the coin metadata fields highlighted the lack of information describing the entire digital collection; missing fields had to be included while some others had to be split, such as, for instance, the coordinates of the item's current location.

As with the coins, the aggregation of the inscription collection started from the metadata extraction used to describe its items. In this case, a complex metadata schema was used, created after an assessment of the available metadata in the field and the discovery that descriptions of all possible components of an inscription were lacking (Vassallo *et al.* [2013](#) 79–82). An analysis of the metadata fields and the comparison with the ARIADNE mandatory criteria was also undertaken. Some cleaning and adjustment of the metadata fields were also necessary in this case (e.g. the inclusion of some fields and merging of subsets).

Enrichment

Enrichment was undertaken on other aspects of the data in order to achieve a higher level of interoperability. In this respect the work continued in parallel with other mandatory steps of the integration procedure, specifically with the enrichment of the metadata for both collections through the use of gazetteers and vocabularies adopted by ARIADNE for standardisation and the use of the relative tools: [PeriodO](#) for time periods and the [Getty Art & Architecture Thesaurus \(AAT\)](#) for subjects.

Chronology standardisation

The first activity consisted of identifying the period(s) covered by the collection in the PeriodO client browser or, where these were lacking, creating any additional ones needed. For both collections, the existence of already published periods in PeriodO corresponding to their chronology facilitated the standardisation process. The exact dates of the Cypriot Frankish and Venetian periods in the gazetteer for the Cypriot medieval coin collection allowed direct matching of the periods of the aggregated collection with those defined in PeriodO. The work dedicated to the standardisation of the epigraphic corpus chronology was a bit more time-consuming but similarly



successful. Indeed, other users had already created chronological periods covering the inscriptions corpus for prehistoric and historic times in Cyprus.

Terminology standardisation

Beyond the time periods, another step had to be taken for terminology standardisation. An open tool integrated into the ARIADNE platform helped us undertake this work. The Vocabulary Matching Tool, developed as part of ARIADNE, supports the creation of mappings from locally used terms or concepts to the AAT (Binding and Tudhope [2016](#); Binding *et al.* [2018](#)). The application presents an editable table of currently derived matches with a direct AAT lookup facility to make more informed mapping decisions. The set of mappings created can be exported to JSON or delimited text (CSV) format for use in other applications.

Therefore, the first step was to extract the terminology used in the Cypriot datasets. In the case of the coin collection, the work was facilitated by the existence of a list of terms (e.g., coin types, mints) created for the DIOPTRA module. All the available terms were inserted into the tool and mapped with the related concept present in the AAT and a JSON file was produced. During the operation, a lack of detailed or precise terms in the AAT for numismatics was noted. This led to a collaboration with other ARIADNE numismatists, particularly with the [Nomisma.org](#) group. The effort is aimed at the homogenisation and standardisation of terminologies used in numismatics. Specifically, this work has been carried out on the terms for medieval coins, mint names and locations, and the issuers under which the coinage was minted. In the case of the inscriptions, the work of standardisation of terms was complicated by the lack of terminology. Consequently, identification and extraction of terms from the dataset were needed before vocabulary matching.

Ontological harmonisation

The next step consisted of the final mapping of the enriched metadata to [CRMtex](#), the CIDOC CRM extension selected by the ARIADNE community for the harmonisation of the datasets and their publication in the portal. CRMtex is devoted to the description of textual entities and is an ideal tool for integrating this kind of information in a wider knowledge graph such as that built by ARIADNE. It comprises a complete application profile that can be used for the modelling of every aspect of texts and their relationship with the monuments, archaeological and artistic objects on which they appear; it is also perfectly harmonised with the ARIADNE AO-Cat ontology (Felicetti and Murano [2021](#); [2022](#); Felicetti *et al.* [2023](#)). CRMtex is, in fact, able to distinguish the two levels of the inscription and its carrier by using the CIDOC CRM class to describe physical objects such as coins, vases and monuments bearing written texts, and by providing specific entities to model the texts themselves and all their individual features. The separation between the physical carrier and text also allows the coherent descriptions of scenarios in which the same object bears two or more texts, as in the case of the inscriptions shown on the obverse and reverse sides of a coin, and to assign different typologies and origin (i.e., production events) to them (Figure 7).



Beyond offering tools for structurally describing textual entities, CRMtex also includes classes and properties for representing the languages, scripts, and alphabets used in writing such entities. Additionally, it models the typical operations performed by scholars when studying texts. The most recent version of the model introduces new classes to describe events of reading, deciphering and understanding, as well as translating and transcribing texts. Being a CIDOC CRM extension, these entities are integrated flawlessly into the CIDOC CRM ecosystem and are fully interoperable with the AO-Cat. The model also offers specific entities for the investigation of defined portions of text and of the interconnections existing between a text and its parts, for instance, text segments, columns, sections and paragraphs, but also single words or letters. This makes it possible to link specific production activities to each individual segment, or specific destruction events, as in the case of letters or words damaged or worn out owing to natural deterioration or human interventions, and to model observations about the condition and state of support, text and its parts, as shown in Figure 8.

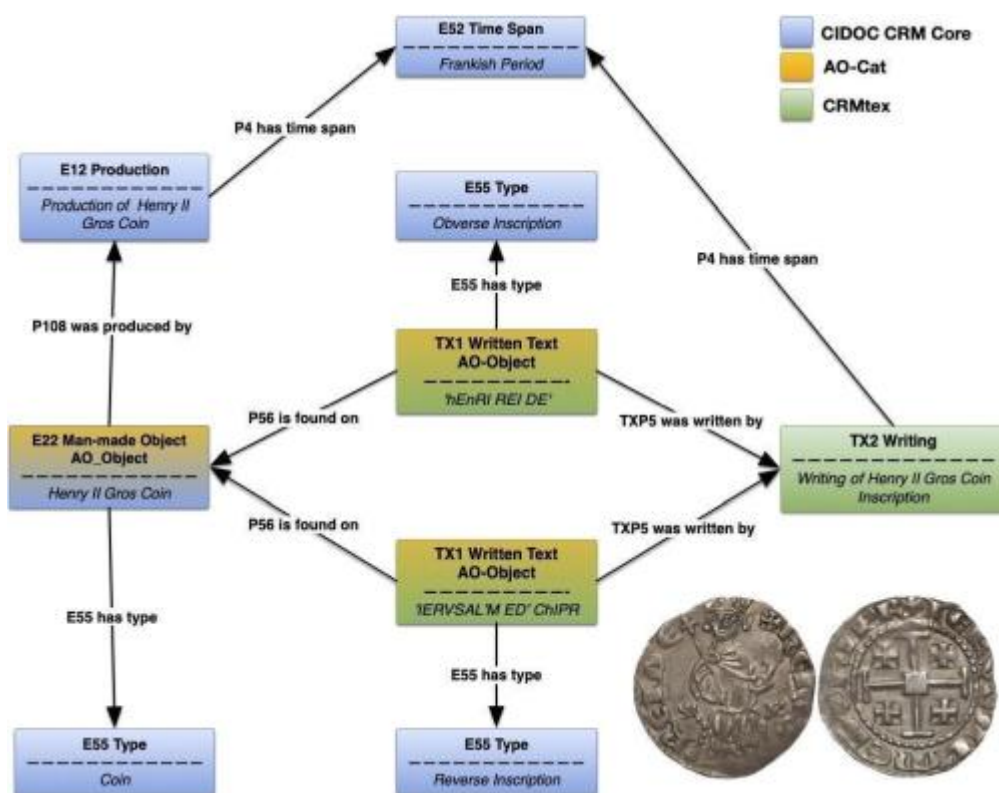


Figure 7: CRMtex, CIDOC CRM and AO-Cat modelling of a medieval coin bearing inscriptions on the obverse and reverse sides

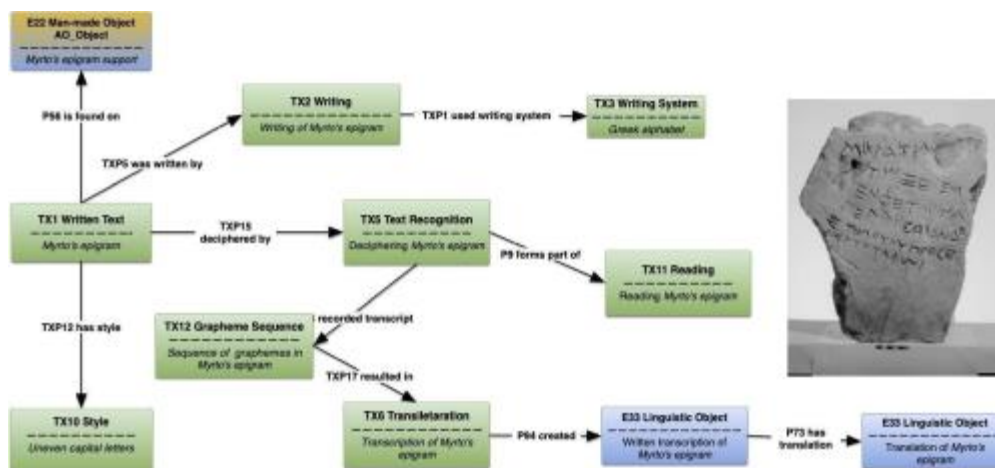


Figure 8: CRMtex, CIDOC CRM and AO-Cat modelling of the various investigation operations (i.e. reading, understanding, transliterating and translating) of an ancient Greek inscription from Cyprus

Eventually, the metadata mapping activity was carried out through the use of the X3ML toolkit. During this phase, some further enrichment was carried out, such as the addition of subject fields to the mapping for describing the collections, as well as checking and fine-tuning the results before final publication.

Publication, semantic browsing and search

The final step was the publication of the metadata in the portal to allow search and data retrieval. First, publication in the staging portal provided an important step for data quality control. Indeed, a security mistake related to the geographical coordinates assigned to the collection during the metadata enrichment, was realised and publication stopped. For collections subject to sensitive information that either should not be published or have to be constantly monitored this can be an important factor (Figure 9).



Figure 9: Publication of the Cypriot medieval coins collection in the ARIADNE Portal

The ARIADNEplus portal offers a browsing facility, allowing the visual exploration of the ARIADNEplus Knowledge Base according to the primary semantic categories defined in AO-Cat at the level of AO_Collections and



AO_Individual_Data_Resources. However, in order to assess the ARIADNEplus ontology, the ARIADNEplus Knowledge Base needed to be exploited at a deeper item-level according to the extra semantic categories defined by the application profiles for the archaeological subdomains. Therefore a demonstrator of a powerful user-friendly semantic searching interface for numismatics and epigraphy was implemented, the main characteristic of which is the support for a flexible exploratory search across the semantic graph, whereby users can transparently construct their semantic queries being guided by the system and based on the knowledge of intermediate query results.

The search interface was built on the [ResearchSpace](#) platform, an open-source collaborative environment for humanities and cultural heritage research using knowledge representation and Semantic Web technologies, developed by the British Museum. An important component of the platform is the Semantic Search, which applies the idea of Fundamental Categories and Relationships (FCs/FRs) to create a semantic query builder. In short, an FC, in a simple implementation, can be considered as a class of entities (usually of a high level). FRs are generic semantic relations among the FCs. So, the key process is the definition of proper FRs that associate two FCs, as shortcut relations that hide the complexity of the corresponding graph patterns that connect the instances of these categories. Like all semantic relations, FRs are directed links from the source-FC to the target-FC. Based on this idea, the Semantic Search interface allows for the definition of categories and relations provided to users for building their semantic queries (Tzobanaki and Doerr [2012](#)). In the ARIADNEplus Semantic Search demonstrator, two searchable categories were provided, *Coin* and *Inscription* and their relations to other categories such as Material, Appellation, Terminology, Place, Actor, Document, Intention and Attribute (Table 1) were described. The list of relations is by no means complete but it is indicative of what can be implemented.



Table 1: ARIADNEplus Semantic Search categories and relations

Coin	Inscription	
consists of		Material
has obverse inscription has reverse inscription		Inscription
has name has title		Appellation
has denomination		Terminology
was produced at	refers to	Place
was produced by issued by published by	was written by refers to read by	Actor
is referred to by		Document
	was intended for purpose was intended for event	Intention
	has language has type is encoded by	Attribute

The basic event flow for building queries, through the UI, is:

1. User: Selects a category to search for instances. This is considered the source-Category of a Relation.
2. System: Shows the user a set of all the associated target-Categories, regarding the previously selected source-Category.
3. User: Selects a category. This is considered the target-Category of the available Relation(s).
4. System: Offers the user a set of all available Relations that can associate the source and the target-Category.
5. User: Selects a Relation.



6. System: Asks about either of the following answers, in order to complete the query statement:
- i. which instance of the target-Category is preferred in the search process. The user at this point has to select a particular resource, found by one of the following options provided by the system:
 - a type-ahead mechanism for any type of labelled resources;
 - a tree selector mechanism, suitable for concepts and place-names hierarchies;
 - a map-selector mechanism, suitable for places with known coordinates;
 - a calendar mechanism for date values and timespans;
- or,
- ii. how the target-Category can be further related to other categories and repeat steps B to F using the target-Category as source for another relation. NB: This step is not currently activated in the demonstrator.

During the query-building process described above, each time a new query statement is constructed and added to the main query, the latter is executed and the result is shown to the user. Then, options are offered for filtering the result according to the relations of the retrieved instances provided by the system. At this point, the user can view the number of retrieved instances grouped under a particular value of each relation, and make a selection of which cases are to be filtered in the final result. The two use cases implemented with this demonstrator are presented to assess the CRM extensions with epigraphic data. The experimental semantic search demonstrator integrated epigraphic data from Cyl (Table 2) and allows the formulation of queries across the datasets.

Table 2: Collections with epigraphic data

Provider	Collection
The Cyprus Institute	Cypriot Medieval Coins History and Culture
The Cyprus Institute	Archaia Kypriaki Grammateia, a corpus of Ancient Greek texts

The datasets are modelled using AO-Cat and CRMtex (Inscriptions, marks and graffiti Application Profile, CRMtex v1.1). In modelling inscription information it was possible to verify that CRMtex, AO-Cat and CIDOC CRM are fully aligned and the formulation of integrated item-level queries is possible. Figure 10 presents the formulation of the query 'Find inscriptions written by Nicocles', while Figure 11 displays the details of a particular inscription.

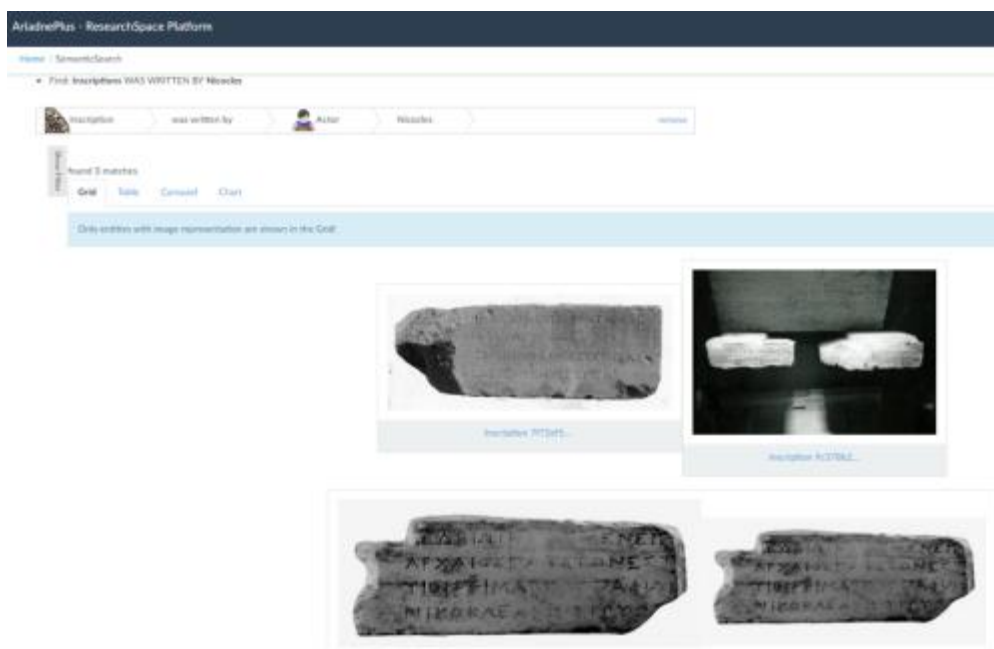


Figure 10: Exploring epigraphs in the Research Space platform

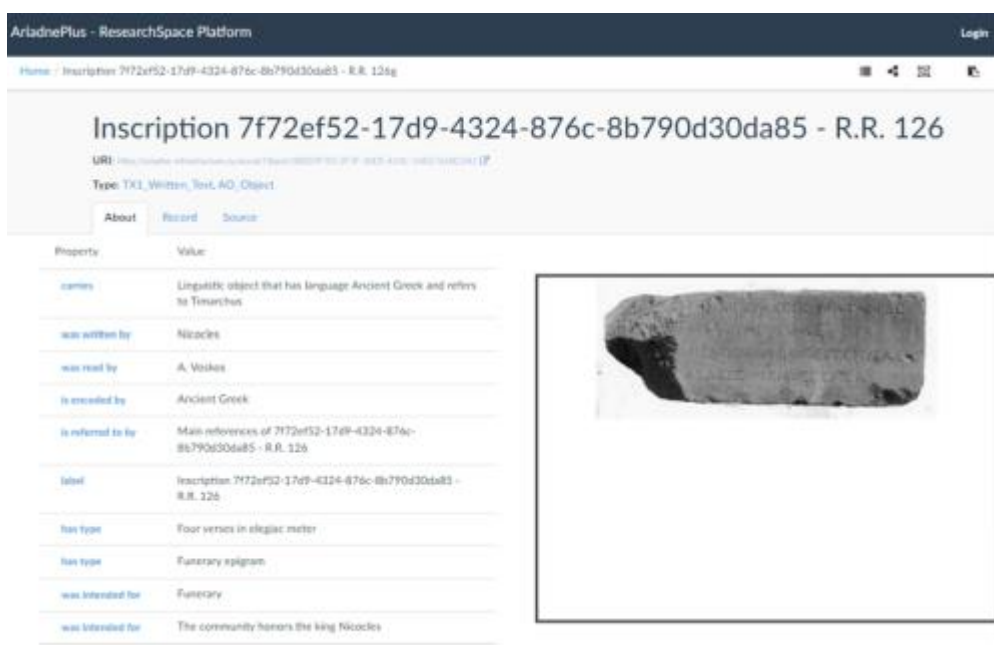


Figure 11: Display of a specific inscription

Indicative queries were run on <https://graphdb-test.ariadne.d4science.org/sparql> which is the SPARQL endpoint of the ARIADNEplus Testing Knowledge Base. At the time of the experiments, some of the epigraphic data was not yet loaded on the official Knowledge Base. Table 3 presents the queries, the respective SPARQL and the results.



Namespaces used:

- PREFIX rdf: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>>
- PREFIX rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>
- PREFIX skos: <<http://www.w3.org/2004/02/skos/core#>>
- PREFIX aocat: <<https://www.ariadne-infrastructure.eu/resource/ao/cat/1.1/>>
- PREFIX crm: <<http://www.cidoc-crm.org/cidoc-crm/>>
- PREFIX crmtex: <<http://www.cidoc-crm.org/cidoc-crm/crmtex/>>
- PREFIX nomisma: <<http://nomisma.org/ontology/>>

Table 3: Indicative inscription queries

Query	SPARQL	No. of results / Publishers
Get a list of archaeological objects bearing an inscription	<pre>SELECT DISTINCT ?object ?objectL ?inscr ?inscrL WHERE { { ?resource rdfs:label ?label; aocat:is_about ?object. {?object crm:P56_bears_feature ?inscr; rdfs:label ?objectL. ?inscr rdfs:label ?inscrL. } UNION {?object a crmtex:TX1_Written_Text; rdfs:label ?objectL.} } } }</pre>	253 on staging 220 from Cyl coins 33 from Cyl Inscriptions
Get a list of inscriptions written on 'White marble' objects	<pre>SELECT DISTINCT ?object ?objectL WHERE { { ?object a crmtex:TX1_Written_Text. ?object rdfs:label ?objectL. ?object crm:P56i_is_found_on ?support. ?support crm:P45_consists_of ?material. ?material skos:prefLabel "White marble"@en. } }</pre>	14 from Cyl Inscriptions
Get a list of languages and scripts used in the inscriptions	<pre>SELECT DISTINCT ?lnote ?snote WHERE { { ?object a crmtex:TX1_Written_Text; rdfs:label ?objectL; crm:P128_carries ?lo. ?lo crm:P72_has_language ?lang. ?lang crm:P3_has_note ?lnote. ?object crmtex:TXP9_is_encoded_by ?script. ?script crm:P3_has_note ?snote. } }</pre>	12 from Cyl Inscriptions



Get a list of scholars who have been involved in the study of inscriptions	<pre>SELECT DISTINCT ?actor ?actorL WHERE { { ?object a crmtex:TX1_Written_Text; rdfs:label ?objectL; crmtex:TXP10i_was_read_by ?reading. ?reading crm:P14_carried_out_by ?actor. ?actor rdfs:label ?actorL. }}</pre>	1 from Cyl Inscriptions
--	---	----------------------------

Find all inscriptions that have a transcription	<pre>SELECT DISTINCT ?transcr ?transcrL WHERE { { ?object a crmtex:TX1_Written_Text; rdfs:label ?objectL; crmtex:TXP10i_was_read_by ?reading. ?reading crmtex:TXP3i_is_rendered_by ?transcr. #?transcr rdfs:label ?transcrL. }}</pre>	0
---	---	---

List all the places mentioned in an inscription	<pre>SELECT DISTINCT ?object ?placeL WHERE { { ?object a crmtex:TX1_Written_Text; crm:P128_carries ?lo. ?lo crm:P67_refers_to ?place. ?place a crm:E53_Place. ?place rdfs:label ?placeL. }}</pre>	28 from Cyl Inscriptions
---	---	-----------------------------

5. Conclusion

The unusual nature and character of the collections described here and their chronological and geographical coverage provided the motivation to integrate them into the ARIADNE infrastructure. Often, some chronological periods or central geographical areas attract much more attention than others, while peripheral regions or more recent times are less considered and studied. The work of aggregation carried out by ARIADNE is very important to attract the interest of the community and the general public to less-known or less-studied geographic areas and chronological periods. Indeed, the possibility through the portal to discover, access and share Cypriot archaeological collections helps to make them visible and to encourage research (and even reuse) both within and across regional and national borders. This is particularly interesting given the historical events that occurred in Cyprus and had an impact on its archaeology, such as the export, smuggling and the dispersion of Cypriot archaeological heritage across the world. Integration in the infrastructure allows the global sharing of Cypriot datasets and, to a certain extent, their re-unification providing, for instance, information on the different locations of the Cypriot archaeological objects, the interoperability between the resources, and their digital preservation.



AO-Cat, CRMtex and the other ontological tools adopted in ARIADNEplus have demonstrated their potential in describing the different semantic nuances of the cultural objects stored in both archives. The ability to fully describe and then relate physical and conceptual objects and to make linguistic considerations for the texts present on the various objects has in fact allowed a faithful representation of their nature and their features in a formal language that makes them easy to query and integrate with other similar information.

Technological tools such as the mapping tools of schemas and vocabularies and the facilities for data collection and grouping, installed and configured as part of the ARIADNE infrastructure, have also proved to be able to perform all the operations necessary for the standardisation, ingestion, enrichment and publication of information relating to epigraphs and coins, and for the integration of these data into the general semantic graph of archaeological data set up by the project.

The presence of a defined aggregation pipeline, the possibility to check all the steps, and the presence of the mandatory fields assisted in the provision of exact information and a correct final product. Moreover, the pipeline and the system in general guarantee updates and this is an important step not only for the checking phase but also in terms of future updates or expansion of the collection with new digital resources. The standard encoding of this data, based on CIDOC CRM, also makes the same information reusable in other contexts where this family of ontologies is used, in full compliance with the FAIR principles.

Bibliography

Avgousti, A., Nikolaidou, A. and Georgiou, R.E. 2017 'OpeNumisma: a software platform managing numismatic collections with a particular focus on Reflectance Transformation Imaging', *Code4Lib Journal* **37**. <https://journal.code4lib.org/articles/12627>

Bardi, A., Binding, C., Felicetti, A., Meghini, C., Richards, J., Theodoridou, M. and Kritsotakis, V. 2022 *ARIADNEplus Data Aggregation Pipeline: User Guide (2.4)*, Zenodo. <https://doi.org/10.5281/zenodo.8060925>

Binding, C. and Tudhope, D. 2016 'Improving interoperability using vocabulary linked data', *International Journal on Digital Libraries* **17**(1), 5-21. <https://doi.org/10.1007/s00799-015-0166-y>

Binding, C., Tudhope, D. and Vlachidis, A. 2018 'A study of semantic integration across archaeological data and reports in different languages', *Journal of Information Science* **45**(3), 364-86. <https://doi.org/10.1177/0165551518789874>

Damnjanovic, U., Vassallo, V. and Hermon, S. 2017 'Integration of multimedia collections and tools for interaction with digital content. The case study of the Archaia Kypriaki Grammateia Digital Corpus' in S. Silvia Orlandi, R. Santucci, F. Mambrini and P.M. Liuzzo (eds) *Digital and Traditional Epigraphy in Context:*



Proceedings of the EAGLE 2016 International Conference, Collana Convegni 36, Sapienza Università Editrice. 247-59.

Felicetti, A. and Murano, F. 2021 'Ce qui est écrit et ce qui est parlé. CRMtex for modelling textual entities on the Semantic Web', *Semantic Web Journal* **12**(2), 169-80. <https://doi.org/10.3233/SW-200418>

Felicetti, A. and Murano, F. 2022 'Semantic modelling of textual entities: the CRMtex model and the ontological description of ancient texts', *Umanistica Digitale* **11**, 163-75. <https://doi.org/10.6092/issn.2532-8816/13674>

Felicetti, A., Meghini, C., Richards, J. and Theodoridou, M. 2023 *The AO-Cat Ontology*. 1.2. <https://doi.org/10.5281/zenodo.7818375>

Hawkins, A. 2022 'Archives, linked data and the digital humanities: increasing access to digitised and born-digital archives via the semantic web', *Archival Science* **22**, 319-44. <https://doi.org/10.1007/s10502-021-09381-0>

Hollander, H., Morselli, F., Uiterwaal, F., Admiraal, F., Trippel, T. and Di Giorgio, S. 2018 *PARTHENOS Guidelines to FAIRify Data Management and make Data Reusable*. <https://doi.org/10.5281/zenodo.2668479>

Jaillant, L. 2022 'How can we make born-digital and digitised archives more accessible? Identifying obstacles and solutions', *Archival Science* **22**, 417-36. <https://doi.org/10.1007/s10502-022-09390-7>

Meghini, C., Scopigno, R., Richards, J., Wright, H., Geser, G., Cuy, S., Fihn, J., Fanini, B., Hollander, H., Niccolucci, F., Felicetti, A., Ronzino, P., Nurra F., Papatheodorou, C., Gavriliis, D., Theodoridou, M., Doerr, M., Tudhope, D., Binding, C. and Vlachidis, A. 2017 'ARIADNE: a research infrastructure for archaeology', *Journal on Computing and Cultural Heritage* **10**(3), 1-27. <https://doi.org/10.1145/3064527>

Palma, G., Corsini, M., Cignoni, P., Scopigno, R. and Mudge, M. 2010 'Dynamic shading enhancement for reflectance transformation imaging', *ACM Journal on Computing and Cultural Heritage* **3**(2), 1-20. <https://doi.org/10.1145/1841317.1841321>

Pitzalis, D., Christophorou, E., Kyriacou, N., Georgiadou, A. and Niccolucci, F. 2012 'Building scholar e-communities using a semantically aware framework: Archaia Kypriaki Grammateia' in D. Arnold, J. Kaminski, F. Niccolucci and A. Stork (eds) *VAST 12. The 13th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*, 89-95. <http://dx.doi.org/10.2312/VAST/VAST12/089-095>

Richards, J.D. 2023 'Joined up thinking: aggregating archaeological datasets at an international scale', *Internet Archaeology* **64**. <https://doi.org/10.11141/ia.64.3>

Richards, J.D., Jakobsson, U., Novák, D., Štular, B. and Wright, H. 2021 'Digital archiving in archaeology: the state of the art. Introduction', *Internet Archaeology* **58**. <https://doi.org/10.11141/ia.58.23>



Richards, J., Felicetti, A., Meghini, C. and Theodoridou, M. 2022 *D4.4 - Final Report on Ontology Implementation*. <https://doi.org/10.5281/zenodo.7636720>

Ronzino, P., Niccolucci, F. and Hermon, S. 2012 'A metadata schema for cultural heritage documentation', *Electronic Imaging & the Visual Arts*, EVA 2012 Florence, 9-11 May 2012. 36-41. <https://media.fupress.com/files/pdf/24/2405/5189>

Stanley-Price, N. 2001 'The Ottoman Law on antiquities (1874) and the founding of the Cyprus Museum' in V. Tatton-Brown (ed) *Cyprus in the 19th Century A.D. Fact, Fancy and Fiction*, Papers of the 22nd British Museum Classical Colloquium. Oxford: Oxbow Books. 267-75.

Tzompanaki, K. and Doerr, M. 2012 'A new framework for querying semantic networks', *Museums and the Web 2012: the International Conference for Culture and Heritage on-line*, San Diego, CA, USA, April 11-14. https://www.museumsandtheweb.com/mw2012/papers/a_new_framework_for_querying_semantic_networks

Vassallo, V., Christophorou, E., Hermon, S. and Niccolucci, F. 2013 'Revealing cross-disciplinary information through formal knowledge representation - A proposed Metadata for ancient Cypriot inscriptions' in A. C. Addison, L. De Luca, G. Guidi, S. Pescarin (eds) *Digital Heritage International Congress (DigitalHeritage)*, IEEE. 79-82. <https://doi.org/10.1109/DigitalHeritage.2013.6744732>

Vassallo, V. and Felicetti, A. 2021 'Towards an ontological cross-disciplinary solution for multidisciplinary data: VI-SEEM data management and the FAIR principles', *International Journal on Digital Libraries* **22**, 297-307. <https://doi.org/10.1007/s00799-020-00285-5>

Vassallo, V., Nunziata, L., Makri, M., Georgiadou, A.S. and Hermon, S. 2023 'The state of the art of digital archiving for archaeology in Cyprus', *Internet Archaeology* **63**. <https://doi.org/10.11141/ia.63.5>

Voskos, A., Michaelides, K. and Taifacos, I.G. 1995 *Αρχαία Κυπριακή Γραμματεία*, Stavrou P. (ed), Leventis Foundation.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva, Bonino, Santos, L., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., Van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenberg, P., Wolstencroft, K., Zhao, J. and Mons, B. 2016 'The FAIR guiding principles for scientific data management and stewardship', *Scientific Data* **3**, 160018. <https://doi.org/10.1038/sdata.2016.18>