D4.1 Initial report on dataset integration

Final Version

8 June 2020

Grant Agreement number:	**823914**
Project acronym:	**ARIADNEplus**
Project title:	**Advanced Research Infrastructure for Archaeological Dataset Networking in Europe - plus**
Funding Scheme:	**H2020-INFRAIA-2018-1**
Project co-ordinator name, Title and Organisation:	**Prof. Franco Niccolucci, PIN Scrl - Polo Universitario "Città di Prato"**
Tel:	**+39 0574 602578**
E-mail:	**franco.niccolucci@pin.unifi.it**
Project website address:	**www.ariadne-infrastructure.eu**

ARIADNEplus is a project funded by the European Commission under the Horizon 2020 Programme, contract no. H2020-INFRAIA-2018-1-823914. The views and opinions expressed in this presentation are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

The research leading to these results has received funding from the European Community's Horizon 2020 Programme (H2020-INFRAIA-2018-1) under grant agreement n° 823914.

Author: **Achille Felicetti, PIN**

Contributing partners: **Maria Theodoridou, FORTH**

Julian Richards, UoY - ADS

Carlo Meghini, CNR-ISTI

Ulf Jakobsson, SND

Harry Dimitropoulos, ATHENA-RC

Holly Wright, UoY ADS (quality control edit)

Table of Contents

1. Executive Summary

This deliverable reports the work done within WP4, comprising Tasks 4.1, 4.2, 4.3, 4.4 and 4.5 during the initial 18 months, i.e. period 1 of the project, assessing it and planning the related activities for the second period, i.e. months 19-36.

The document is structured according to the activities carried out by the work package. After an overall presentation of the main aims of WP4 in Chapter 2, Chapter 3 provides an overview of the conceptual model being developed in ARIADNEplus and its various components. It focuses on the new AO-Cat model, designed to describe the datasets in the Catalogue and improves the previous ACDM model developed in ARIADNE. The full AO-Cat ontology is provided in the Appendix of this deliverable. The chapter also presents the fundamental categories defined to classify the information in the Catalogue and make it easily available on the new ARIADNEplus Portal, according to the recommendations of the FAIR principles which inspired the model. An overview of the application profiles under development in WP14 along with a review of the compatible models already in use by some partners.

Chapter 4 presents the tools developed and implemented in ARIADNEplus to facilitate the encoding process of metadata according to the AO-Cat model and to assist content providers in all phases of their ingestion, mapping, transformation, enrichment and publication activities. Of great importance in this sense is the 3M Mapping Tool developed by FORTH, which allows users to define in detail the correspondences between the legacy metadata schemas and the entities of the model, in order to implement an optimal level of integration. The Fast Cat tool, designed for the rapid acquisition of information directly in the AO-Cat format, is also offered to partners who do not use any format for their metadata or who have a limited amount of data to be provided.

Chapter 5 presents the Helpdesk, a collaborative service provided by the ARIADNEplus platform to assist content providers in all phases of the data contribution and to provide assistance at every stage of the process, from preparation to the definition of mappings, up to the fine-tuning the data harvesting and data acquisition mechanisms in the ARIADNEplus infrastructure. The service is based on the ticketing system and offers efficient interaction with the special team of experts set up to provide all the necessary information to foster the process.

Chapter 6 documents the status of the integration, shows how the partners are adapting their data to the ARIADNEplus standards and the priorities defined for ingestion, according to the progress of these operations. The encoding, enrichment and standardisation work also relies on the use of the various vocabularies adopted by ARIADNEplus (e.g., the Getty AAT for subjects and PeriodO for time periods) and documented in the deliverable D5.2. Particular attention is paid to the mapping operations and a detailed analysis of the progress made on these activities is provided for each partner and each discipline listed in Task 4.4.

Chapter 7 presents the activities aimed at linking the ARIADNEplus Data Infrastructure with repositories of scientific publications, exploiting, in particular, OpenAire and its open access digital archives and the links to individual journals such as Internet Archaeology or A&C. The chapter also

describes the use of the ARIADNEplus text mining service (Task 17.4) to improve the metadata for the textual resources.

The conclusions and an evaluation of the activities carried out in the first 18 months of the project are presented in chapter 8. The same section presents the strategies proposed for future work and for the completion of WP4 activities.

2. Introduction and Objectives

This deliverable reports the work done within WP4, comprising Tasks 4.1, 4.2, 4.3, 4.4 and 4.5 during the initial 18 months, i.e. period 1 of the project, assessing it and planning the related activities for the second period, i.e. months 19-36.

The main objective of WP4 is to foster the integration of datasets from the ARIADNEplus archaeological research community. ARIADNEplus aims to integrate the best and most important research infrastructures within the archaeological domain and provide services for effective access and re-use of data.

ARIADNEplus is defining a data model built around a standard that is progressively becoming universally accepted in the archaeological domain: the CIDOC Conceptual Reference Model (CRM)[1]. CIDOC CRM is a formal ontology in the form of a semantic model, i.e. it declares each attribute value as an *a priori* independent entity in the universe of discourse and connects such values symmetrically by directed links/arcs/properties. For its peculiarities, CIDOC CRM has rapidly conquered a large portion of the Cultural Heritage domain, as witnessed by its influence on the Europeana Data Model. The ARIADNE Data Model inherits the CIDOC CRM model and philosophy to provide the community with a powerful and efficient integration tool, capable of capturing the meaning of legacy archives offered by content providers and expressing their data in a unique and coherent way.

WP4 activities are focused on facilitating the adoption of the ARIADNEplus Data Model within the consortium by means of recommendations, advice and tools aimed to simplify the conversion of legacy data and their encoding in a standard format. In particular, it advises users about the construction of data schemas appropriate to their needs, the choice of convenient data management software and, if necessary, the conversion of pre-existing databases to the new ARIADNEplus infrastructure. Support to the technical design and definition of national or thematic archaeological data infrastructures are also envisaged, through the collaboration of local partners.

According to these objectives, WP4 deploys a series of strategies aimed at implementing and strengthening the integration and interoperability, that mainly focus on:

1. Create and offer services to support users in contributing to the ARIADNEplus integration objective.
2. Provide guidelines and tools to integrate datasets and data in the ARIADNEplus European Archaeological Data Infrastructure.
3. Support users in mapping and/or conversion of their archaeological metadata and datasets to the ARIADNEplus standards.
4. Facilitate the integration of archaeological data into the ARIADNEplus Cloud.

[1] http://www.cidoc-crm.org

Other WP4 tasks concern the promotion of existing standards and the use of application profiles set up in WP14 within the infrastructure. It also fosters the integration of ARIADNEplus with existing digital libraries, notably OpenAire.

In ARIADNEplus, standardisation is defined bottom-up and comprises the following steps, as shown in Figure 1:

1. Definition of the necessary application profiles through the collaboration of selected users, i.e. archaeologists with expertise in the different domains providing data for integration, and experts in knowledge engineering. This will produce draft proposals to be checked against perceived user needs and then formalised.
2. Mappings are created (with the mapping service X3ML) and tested on samples from the various involved datasets. Feedback will suggest amendments. This will produce the necessary mapping sets.
3. Vocabularies and gazetteers are selected and improved/adapted.
4. When the mapping is stable, the mapping/conversion/ingestion process starts.

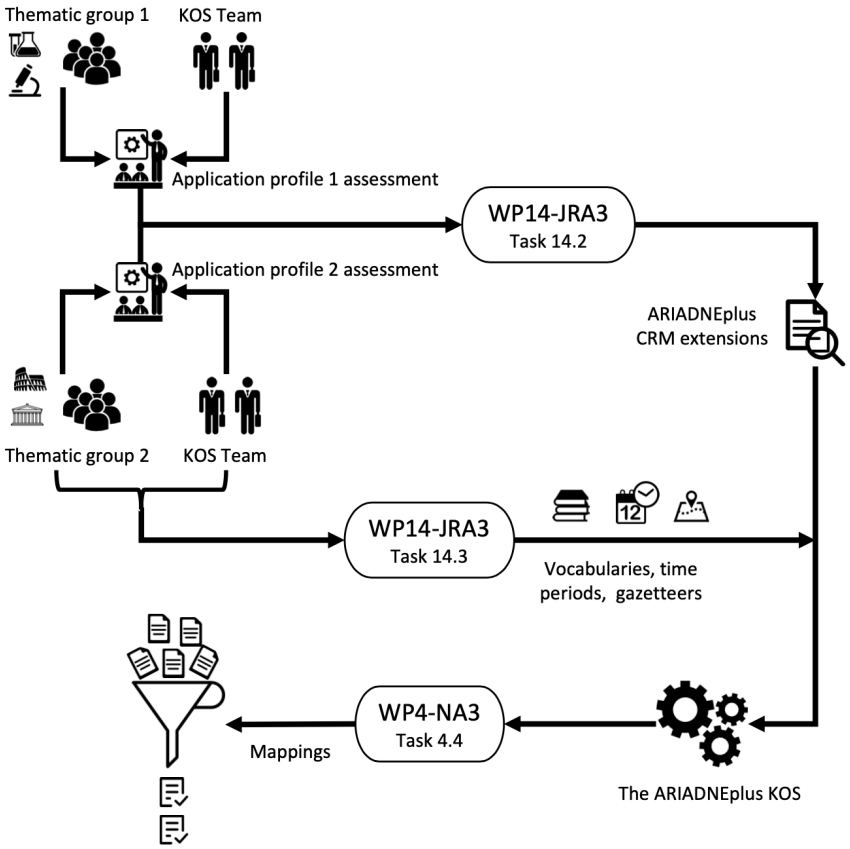


Figure 1: The ARIADNEplus knowledge organization system (KOS) set-up procedure.

WP4 mainly operates in the context of point 2, with particular regard to the mapping and the alignment of the data models in use in the various disciplines for which ARIADNEplus provides coverage. Application fields include excavation data, information produced by palaeoanthropology, bioarchaeology, environmental archaeology as well as the results of scientific analyses, such as material sciences and those related to standing structures.

3. The ARIADNEplus Data Model

The ARIADNE Content Cloud (AC for short) is the information repository collecting all the knowledge managed by the ARIADNE infrastructure. The AC is a composite knowledge base that can be understood as consisting of several related parts, namely:

- the ARIADNE Catalogue, giving general knowledge about the resources of the AC; data resources will be represented in the Catalogue at the collection level;
- several sub-knowledge bases, each containing knowledge statements about a specific Archaeological domain.

The ARIADNE Ontology (AO for short) is the ontology providing the terms for making the statements in the AC. The AO is a formal ontology of the resources managed by a research infrastructure, with a special focus on the archaeological domain and on its infrastructure, ARIADNE, developed by the ARIADNEplus project.

Mirroring the structure of the AC, also the AO will be structured into sub-ontologies, namely:

- a Catalogue ontology, named AO-Cat, providing terms for the statements within the ARIADNE Catalogue;
- several application profiles, providing terms for the domain-specific parts of the AC.

The ARIADNEplus Data Model is a language providing syntax and semantics for the AC and the AO. Conceptually, both the AC and the AO will be expressed in OWL 2 DL using the direct semantics for compliance with the standards. Moreover, the AO will be based on consolidated standards in the Cultural Heritage domain, chiefly the CIDOC CRM. Concretely, both the AC and the AO will be expressed in the Resource Description Framework (RDF for brevity) language for knowledge representation, using convenient concrete notations amongst the several existing serialisations (Turtle, XML-based, etc.). Consequently, AO classes and properties are IRIs in the ARIADNE namespace, which is identified by the IRI:

https://www.ariadne-infrastructure.eu/resource/ao/cat/

Class and property IRIs are written as qualified names using no prefix, meaning the ARIADNE namespace IRI is used as a base IRI that applies to all of them. For example, the class of temporal regions is named as http://ariadne-infrastructure.eu/ns/AO_Temporal_Region and in this document is written as AO_Temporal_Region.

3.1 The AO-Cat ontology for ARIADNE Catalogue

AO-Cat derives from the ARIADNE Catalogue Data Model (ACDM for short), employed in the predecessor ARIADNE project to model archaeological resources, and from the PARTHENOS Entities Model (PEM), employed in the PARTHENOS project to model the resources managed by a research infrastructure. In its present version, AO-Cat is a contraction of the ACDM driven by the conceptualisation underlying the PEM. In addition, AO-Cat inherits from the PEM its tight relation to the CRM, which is used to represent any aspect of archaeological resources not covered by the ACDM or the PEM. Those aspects that are not covered by any of the three models mentioned so far, will be dealt with by introducing ad hoc terms in AO-Cat.

The rest of this Section is structured as follows: the requirements informing AO-Cat are firstly outlined, and then a general description of the ontology will follow. The full AO-Cat ontology is provided in the Appendix.

Requirements

There are two basic requirements for the ARIADNEplus Catalogue and more generally for the AC:

1. Enabling cross-border/cross-institution resource discovery, i.e. finding data.
2. Enabling interoperability–across partners, countries, data types, data schemas, i.e. enabling research.

Concerning resource discovery, the Catalogue should support:

- *What* searches, searches based on a subject. For subject, the Getty AAT should be used as the common conceptual backbone, in addition to local reference resources.
- *When* searches, searches based on a temporal period. For temporal periods, the PeriodO vocabulary should be used as it maps periods to absolute dates on a common time scale. In addition, local vocabularies providing region-related period appellations should be represented in the AC and allowed in searches by cultural period e.g. “Iron Age”.
- *Where* searches, searches based on a spatial region. For spatial regions, the World Geodetic System 1984 (commonly abbreviated as WGS84) representation should be used.
- Enhanced map-based searching.
- Enhanced queries, specific to data types.

More search types that should be supported will be defined by a user needs study undertaken by the project.

To enable research, in addition to the above types of search, the AC is expected to store information about digital objects belonging to a large range of data types, forming a heterogeneous set, as data is diverse in character and content. These data types include, but are not restricted to:

- Databases
- Reports
- Finds

- Images
- GIS
- LiDAR data
- Datasets e.g. excavation archives
- Sub-domains, e.g., scientific data; these should be modelled as CRM Application profiles, that is subsets of the CRM that we use for item level integration (cf ARIADNE coins)
- Linguistic resources, such as ontologies and vocabularies relevant to the archaeological domain

In addition, different levels of granularity must be accounted for, supporting collection and item level interoperability, e.g., sites and the individual artefacts they contain; individual dates; cemeteries and the individual graves they contain. Moreover, links to distributed digital and paper resources about the described resources should also be maintained. Finally, archaeological fieldwork events should be accounted for, categorised following standard vocabularies and properly connected to the relevant archaeological resources, as schematised in Figure 2 below, where each event category is annotated with the number of the sub-task of A+ Task 4.4.0 addressing it:

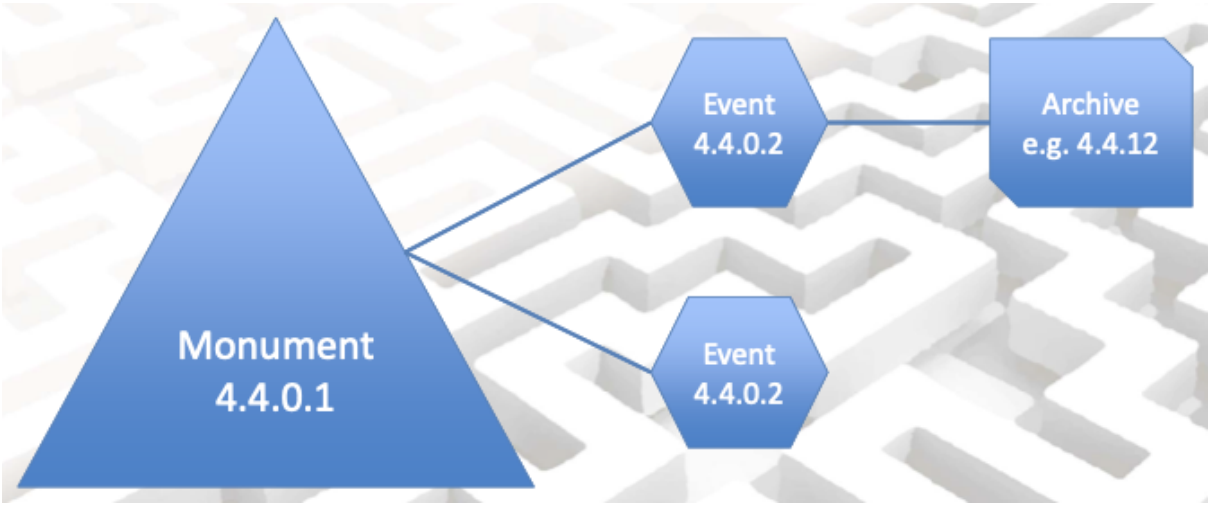


Figure 2. Entities and relationships in the archaeological domain.

The diagram follows the UK MIDAS heritage standard terminology, whereby a 'Monument' is a physical entity where activity took place in the past, possibly over several periods e.g. Stonehenge; whereas an 'Event' normally refers to an archaeological recording event at that monument, such as an excavation or field survey. e.g. "Excavations at Stonehenge by Richard Atkinson 1950-64". In other cases, *e.g.* the "Tiber Valley Survey", a single archaeological event (the survey) may lead to the discovery of multiple new archaeological monuments, or sites.

The services described above should be offered both to humans, via the ARIADNEplus Portal and supporting multilingualism, and to artificial agents, via APIs. Those APIs should provide access to the whole AC (including the Catalogue) as Linked Open Data, thereby allowing other organisations to implement their own portal or service.

3.1.1 Overview of AO-Cat

This Section provides an overview of AO-Cat by showing how the requirements presented in the previous Section have been taken into account to define the main classes of AO-Cat and the properties that have those classes as domains.

The AO-Cat namespace

It has been decided to define a new namespace for the AO-Cat classes and properties as opposed to re-using the namespace of an existing ontology (the CIDOC CRM would have been a natural choice) in order to freely define the classes and properties needed to match the requirements. It may be that the vocabulary of the CRM, including that of the CRM extensions, is adequate for the needs of ARIADNE. In this case AO-Cat will simply be an application profile of the CRM, but this will only be discovered once the AO-Cat is fully developed. Until that time, new vocabularies will be created and its terms are systematically mapped to those of the CRM, so that at any time it is possible to know the relative position of the ontology being developed with respect to the CRM.

AO_Entity

This is the most general class of AO-Cat, all classes being sub-classes of AO_Entity. As such, AO_Entity has as instances all resources that have any role in the ARIADNE infrastructure. AO_Entity is defined for capturing domains or ranges of properties that cover the full ARIADNE information space, such as the range of the is_about property.

There are a few properties that have AO_Entity as domain, and therefore apply to all resources:

- has_type, a type of the resource in any classification system.
- has_title, a title of the resource.
- has_description, a textual description of the resource using natural language.

AO_Entity is the range of property is_about, having AO_Data_Resource (see below) as domain, to express the fact that a data resource may document any entity in ARIADNE.

ARIADNE Infrastructural Resources

The most general infrastructural resource class in AO-Cat is AO_Resource, representing all the digital resources the ARIADNE research infrastructure deals with. The rest of this Section gives some fundamental principles that AO-Cat follows in modelling infrastructural resources, including the sub-classes of AO_Resource.

An ARIADNE resource is a web resource, therefore it is identified in the AC by an HTTP IRI in the ARIADNE domain, which is given by: https://www.ariadne-infrastructure.eu/ The identifier of the resource in this namespace is called the “primary” resource identifier and is computed at aggregation time from the local identifier of the resource.

In addition to the primary identifier, the AC also represents the identifier of the resource in the namespace of the provider. This identifier, which is called the “local” resource identifier is associated to the resource via property has_original_id which has simply character strings as values, to allow

maximum generality. It is understood that the local identifier of a resource is the identifier used by the publisher of the resource, which is the Agent making the resource publicly available (see below).

Finally, the AC also allows representing other identifiers of a resource than the one in the publisher namespace, via the property has_identifier Also this property ranges over character strings for generality. As a consequence, it is not possible to represent any additional knowledge about these other identifiers other than the knowledge that can be inferred from the identifiers themselves.

Any ARIADNE Resource has a type which can be specified in two different ways:

1. one way is via the property has_type, which associates an instance of AO_Entity with a type represented as an AO_Concept, as previously seen;
2. the other way is via the sub-classes of AO_Resource, discussed below.

The former way is preferred when the type of resource does not have any significance other than a tag attached to the resource. The latter is preferred when resources of different types can have different properties.

In addition to the properties inherited from AO_Entity, a number of descriptive properties are defined for an ARIADNE resource:

- was_issued, the date when the record of the resource was firstly acquired.
- was_modified, the date when the record of the resource was lastly modified.
- has_publisher, the agent responsible for making the resource publicly accessible.
- has_contributor, a contributor of a description of the resource to the AC.
- has_creator, a creator of the resource.
- has_owner, the owner of the resource.
- has_responsible, any person who is scientifically responsible of the resource.

These properties have been chosen based on the experience gained in the first ARIADNE project, and are described in detail in the Appendix.

Data Resources

The ARIADNE data space includes data resources belonging to a wide range of data types, because archaeologists typically employ a wide, open-ended range of information technologies in their research, producing data belonging to a correspondingly open-ended range of types. For the moment, AO-Cat treats these data types simply as tags because no specific property is defined for a specific data type. As such, they are represented as instances of class AO_Concept and are associated with instances of AO_Data_Resource using the property has_type.

Data resources are also characterised in terms of the entity they are about. The connection between a data resource and the entity the resource is about is captured by property refers_to:

- refers_to associates a data resource with an entity to which that the resource refers, by making assertions, whether implicitly or explicitly and regardless of the format, about that entity.

As a special case of reference, there is also the property is_about.

- is_about associates a data resource with the primary entity that the data resource documents, including events. Notice that the data resource being documented may be an external resource, identified by an IRI, or DOI, or any other identifier, to capture the requirement that the ARIADNE AC should have links to distributed digital and analogue resources outside ARIADNE.

In turn, is_about is categorised into three sub-properties making it possible to trace the provenance of the association between the data resource and the subject of the entity:

- has_native_subject models the association directly imported from the provider of the data.
- has_derived_subject models the association computed by the ARIADNE aggregator by mapping the native subject to an ATT term, to match the requirement that "the Getty AAT should be used as the common conceptual backbone, in addition to local reference resources".
- has_ARIADNE_subject models the association between the data resource and one of the fundamental ARIADNE categories.

In addition to topicality, required to support What searches, the requirements highlight two important aspects of a data resource:

- the spatial coverage, required to support Where searches;
- the temporal coverage, required to support When searches.

To account for these aspects, AO-Cat introduces the following properties:

- has_spatial_coverage, associating an instance of AO_Data_Resource with a spatial region;
- has_temporal_coverage, associating an instance of AO_Data_Resource with a temporal region.

Spatial and temporal regions are therefore required in AO-Cat, and the classes and properties to model them are introduced in the Appendix.

As for AO_Resource, several descriptive properties are defined for data resources, and as such have AO_Data_Resource as domain. These are:

- has_language, the language of the resource.
- was_created_on, the date upon which the resource was created.
- has_landing_page, the original landing page of the AO_Data_Resource, if any.
- has_access_policy, a statement of access policy for the resource.
- has_access_rights, a statement of access rights on the resource.
- has_extent, the size of the resource.

The requirements highlight the importance of accounting for the different levels of granularity of the data space, therefore AO-Cat distinguishes data resources into two basic types (the complete resource class taxonomy is given in the Figure 3):

- individual data resources, which are atomic in the sense that they are not further decomposable in other resources as far as the ARIADNE infrastructure is concerned (these are represented as instances of class AO_Individual_Data_Resource, a sub-class of AO_Data_Resource);
- collections, which are wholes composed of parts that are data resources themselves, and as such they may be collections; collections are represented as instances of class AO_Collection.

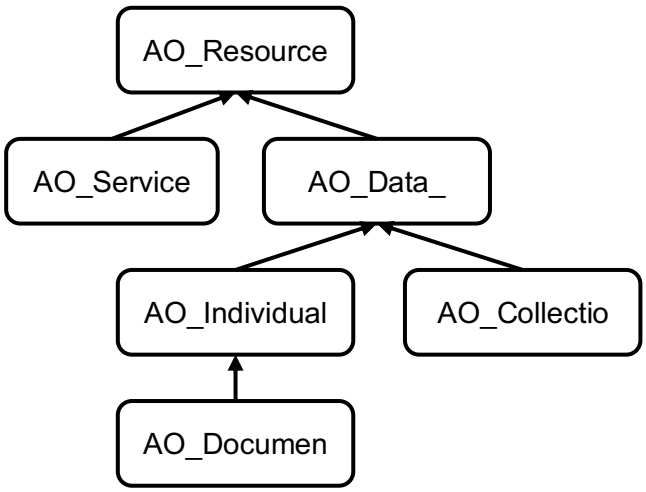


Figure 3. Resources class taxonomy in AO-Cat.

Archaeological records provided to ARIADNE by its members are modelled as individual data resources, while property is_about is used to associate each record with the entities (including events) about which it primarily carries information. A special kind of individual data resources are documents, to accommodate which class AO_Document is created as a subclass of AO_Individual_Data_Resource. In particular, AO_Document is the class of all the documents which which ARIADNE needs to work, such as unpublished fieldwork reports (grey literature), published journal articles, radiocarbon dating resources, etc., whose content may possibly be part of the AC.

The AO-Cat property has_part is introduced to associate a collection with its members.

Services

The class AO_Service has services as instances. Following the PEM definition, a service is “an offer by some actor of their willingness and ability to execute an activity or series of activities upon request”

The descriptive properties defined on services are:

- is_accessible_at, associating the service to an IRI identifying the resource making the service accessible. If the service is a web service, this IRI is the technical access point of the service. Otherwise, it is the IRI of a resource describing how the service can be accessed.
- has_functionality associating the service to a concept giving the functionality of the service.
- has_consumed_media, associating the service to concepts giving the consumed media type.
- has_produced_media, same as before for the media type produced by the service.
- has_consumed_format, associating the service to concepts giving the consumed media format.

- has_produced_format, same as before for the media format produced by the service. A natural candidate vocabulary for the last four properties is the IANA MIME type hierarchy[2].
- has_supported_language, associating the service to concepts giving the supported natural languages. A natural candidate for this property is the language vocabulary established by Lexvo.org.
- has_technical_support, associating the service to an agent providing technical support for the service.

Objects

Objects play an important role in the archaeological domain: much of the activity, e.g., in an excavation, concerns objects: discovery, analysis, classification, dating, and so on. As such, objects form an important category for contextualising the digital resources of the ARIADNE infrastructure and therefore this category needs a place in the AO-Cat.

Class AO_Object is defined for the purpose of classifying within the ARIADNE AC, all physical objects relevant to ARIADNE. This includes three defined properties:

- has_time_interval, to relate an object to a temporal region relevant to the object, such as the period when the object was created.
- has_space_region, to relate an object to a spatial region relevant to the object, such as the place where the object is located.
- was_present_at, to relate an object to an archaeological event relevant to the object, such as the excavation event that led to a discovery of the object.

Concepts

There are many concept spaces that can be used to describe the topics a data or a service resource is about. Generally speaking, these topics are grouped in three main categories in the ARIADNE AC:

- the ARIADNE fundamental categories;
- the terms of the AAT Thesaurus;
- any other term used in the data of some provider.

In principle, each of these categories is expected to evolve. For instance, the ARIADNE fundamental categories currently are

- Site/monument
- Fieldwork
- Fieldwork report
- Scientific analysis
- Date
- Artefact
- Fieldwork archive
- Inscription

[2] https://www.iana.org/assignments/media-types/media-types.xhtml

- Burial

These categories emerged as *ad hoc* pragmatically defined data clusters from the first ARIADNE project, and it is anticipated they will evolve and increase as the WP4 subtasks gather information about data categories commonly held by ARIADNEplus partners. The same applies to the other two categories, if for no other reason because both are outside the control of the ARIADNE Community.

AO-Cat provides the class AO_Concept to represent topics in each category. Moreover, to allow providers to associate the necessary information with each concept, AO_Concept is declared equivalent to skos:concept. As a consequence, all properties defined by SKOS on concepts can be applied to any instance of AO_Concept. For instance,

- the properties skos:broader and skos:narrower can be used to model concept taxonomies;
- the properties skos:broadMatch, skos:closeMatch, skos:exactMatch, skos:narrowMatch, skos:relatedMatch can be used to specify the kind of mapping leading to the instance of the concept;
- the property skos:inScheme can be used to specify the ontology or the terminology from which a concept is derived.

The reason why it has been decided to re-use SKOS (as opposed to defining the above properties in AO-Cat) is that SKOS is now a *de facto* standard, thoroughly tested over many years of application in a wide range of contexts, confirming SKOS will be adequate for concept modelling within ARIADNEplus.

Spatial regions

The generic AO-Cat class for spatial regions is AO_Spatial_Region. Two properties are defined for the instances of this class:

- has_coordinate_system
- has_place_name

Based on the experience of the many content providers in the ARIADNEplus Consortium and of the previous ARIADNE project, four main representations of a spatial region are provided by AO-Cat, each assigned to a specific subclass of AO_Spatial_Region. These subclasses are (see Figure 4 below):

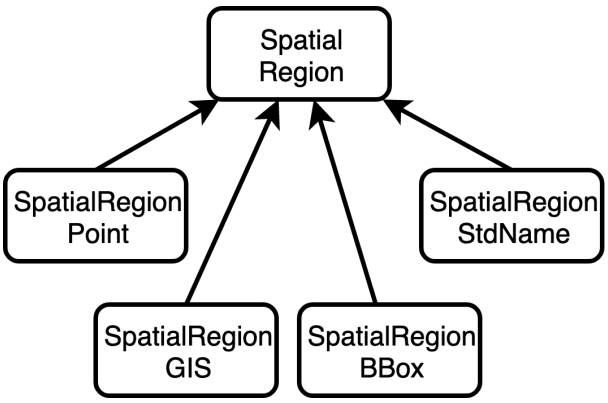


Figure 4. Spatial entities in AO-Cat.

- AO_Spatial_Region_Point, having as instances regions that are points. The properties defined on points are:
    - has_latitude
    - has_longitude
- AO_Spatial_Region_Polygon, having as instances regions that are polygons, represented in some format typically managed by a GIS. AO-Cat does not require a specific notation format for these regions; they are treated simply as abstract objects with a specific property:
    - has_polygonal_representation;

  giving the XML document representing the polygonal region.

- AO_Spatial_Region_BBox, having as instances regions that are bounded boxes, represented by four points giving the vertices of the box. The properties defined on points are:
    - has_bounding_box_min_lat
    - has_bounding_box_min_lon
    - has_bounding_box_max_lat
    - has_bounding_box_max_lon
- AO_Spatial_Region_StdName having as instances regions that are simply identified by a standard name in some vocabulary. The standard name is given by the property:
    - has_place_IRI

Temporal regions

Time is represented and named in many different ways in the archaeological domain, often inextricably associated with space. Fortunately, reference resources have been created in the last decade, which make the task of accounting for time much easier for a research infrastructure such as ARIADNE.

According to the requirements, AO-Cat should support both time points, defined as absolute dates with respect to a reference system, and time intervals, defined as temporal extents having a beginning, an end and a non-zero duration. In addition, time intervals identified by names (*e.g.*, “Neolithic”) should be supported, whether these names are drawn from a standard or a local vocabulary.

For the representation of time points, AO-Cat relies on the “date” data type of the XML Schema type system, written as xsd:date.

For the names and representations of time intervals, AO-Cat relies on the PeriodO gazetteer service. PeriodO allows the definition of a temporal interval as a web resource, associated with a label and a pair of absolute dates giving the earliest start and the latest stop of the interval. The service also allows the clustering of period resources within collections, thus facilitating the exploration of the gazetteer data space.

The AO-Cat class AO_Temporal_Region has time intervals as instances. Each instance can be described using the following four properties:

- has_period, giving the PeriodO time interval as an IRI.

- has_native_period, giving the local identifier of the period, as an instance of AO_Concept.
- from, giving the beginning of the interval as an xsd:date.
- until, giving the end of the interval as an xsd:date.

Events and Activities

Much archaeological research activity concerns field work. Field work results in digital data resources being collected or created for the purpose of documenting the research activity, presenting its results, and communicating synthetic arguements. Digital data resources may also be the result of analyses conducted in scientific laboratories.

In AO-Cat, these research activities are generally modelled as activities, in the sense of "actions intentionally carried out by actors that result in changes of state in the cultural, social, or physical systems documented. This notion includes complex, composite and long-lasting actions such as the building of a settlement or a war, as well as simple, short-lived actions such as the opening of a door" [CIDOC CRM Specs 6.2]. AO-Cat follows the conceptual structure of the CRM and models activities as special kinds of events, where an event is a "change of states in cultural, social or physical systems, regardless of scale, brought about by a series or group of coherent physical, cultural, technological or legal phenomena" [CIDOC CRM Specs 6.2].

Events and activities play an important role in contextualising the data resources held by a research infrastructure, and ARIADNE is no exception. That said, it must be noted not all events and activities involving an ARIADNE resource are equally well documented and require an explicit representation. For instance, the production of a vase found during an archaeological excavation documented in the ARIADNE Content Cloud, is an activity of the remote past about which very little is known. It is common practice not to focus on that activity and emphasise what is known about it directly with the object resulting from the production event. These operations, abstracting the activity and associating the knowledge about the event directly with the resulting object, is sometimes termed "shortcutting"

In sum, AO-Cat provides the classes AO_Event and AO_Activity to explicitly represent the events and the activities (respectively) that are sufficiently documented to be a resource within the ARIADNE research infrastructure; while it provides properties such as has_temporal_region or has_spatial_coverage (already introduced) to represent knowledge about poorly documented activities, and to associate that knowledge directly to data resources.

The following properties are defined on instances of AO_Event and inherited by instances of AO_Activity:

- occurs_in, to associate an event with the spatial region where the event has taken place.
- happens_at, to associate an event with the temporal region when the event has taken place.
- contains_event, to associate a composite event (such as the Trojan War) with any constituent event (such as the duel between Achilles and Hector).

The last property is introduced to represent that a collection, which is a composite data resource, may document a single activity, while the members of the collection may document parts of that activity. These parts are events themselves contained within the activity documented by the collection, and that containment relation is captured by the property contains_event.

Agents

Agents play important roles in the ARIADNE information space: they are responsible for making resources available and publicly accessible; they hold various types of responsibilities for those resources; finally, they carry out activities. For these reasons, AO-Cat defines the class Agent to model entities that can act.

Moreover, from an archaeological perspective it is important to distinguish between two kinds of agents: the person and the organisation: Julian Richards as the individual and researcher and/or excavation director (who might have an ORCID or other id) and University of York as the organisation that was legally responsible for the excavation or its funding, or publication etc. This distinction is captured in AO-Cat by two subclasses of Agent:

- Person, modelling individual humans, and
- Group, modelling "gatherings or organizations of agents that act collectively or in a similar way due to any form of unifying relationship" [CRM Spec 6.2].

The following properties are defined on instances of Agent, regardless of whether the agent is a person or a group:

- has_name, giving the name of the agent.
- has_agent_identifier, giving any identifier of an agent outside of the ARIADNE namespace.
- has_email, giving the email address of the agent.
- has_homepage, giving a web page for the agent.
- has_insitution, giving the institution(s) of the agent.

Modelling the basic pattern

The way AO-Cat deals with a basic pattern in archaeological modelling, such as the one presented in the previous Figure includes:

- the core entities in the Figure are the two events in the center of the Figure. They are modelled as instances of class AO_Event;
- the monument in the big triangle is modelled as an instance of AO_Object and;
- the archive in the small rectangle is an instance of AO_Data_Resource.
- The relations between the monument and each event occurrence is represented by instances of property was_present_at;
- the relation between the archive and the event is represented by an instance of property is_about.

3.2 Application profiles and other compatible data models

Application profiles are ontological extensions defined to describe in detail the specific entities of various disciplines. In ARIADNEplus, partners involved in task 4.4 provide input about the requirements of each sub-discipline, which is then synthesised and verified by domain experts. This leads to the definition of CIDOC CRM compatible application profiles for each sub-discipline, many of

which will be based on CRMarchaeo[3], the extension of CIDOC CRM specifically designed to model archaeological excavation activities and all the related entities. Their release will take place during the second half of the project, as components of the ARIADNE Data Model, since their contribution will be significant especially during item level integration work. Some, however, are already in an advanced design phase, especially those relating to inscriptions and the encoding of scientific data, which are presented below.

Inscriptions

Various activities for the definition of an application profile for inscriptions took place during the first half of the project, and some existing models already in use by some ARIADNEplus partners were examined for this purpose. The most relevant models are CRMtex[4] (by PIN and FORTH) and EPNet[5] (by University of Barcelona).

CRMtex, an ontological model based on CIDOC CRM created to describe ancient texts and other semiotic features appearing on inscriptions, papyri, manuscripts and other media. The model is also designed to describe in a formal way the phenomena related to the production, use, conservation, study and interpretation of textual entities. EPNet is designed to deal with the inscriptions, events and objects connected with the distribution of food in the Roman world and is CIDOC CRM based as well. The full compatibility of these models with the CIDOC CRM ontology and its extensions ensures persistent interoperability of data encoded by means of its entities. This is along with other semantic information produced within cultural heritage and digital humanities, and those within the ARIADNEplus ecosystem, based on the same conceptual reference model. Additionally, the metadata models of many other partners (CyI, SND, CNR) already including fields/classes for the documentation of inscriptions, graffiti, marks, rock art and other similar material, were investigated.

Among the other interesting initiative outside ARIADNEplus, of particular importance is the work done by the EAGLE Project[6]. EAGLE has developed a CIDOC CRM compatible metadata model[7] which represents an excellent (although very complex) example of an application profile designed for a specific purpose (i.e. integrating epigraphic resources). EAGLE has also released a complete set of extremely useful vocabularies[8] for describing epigraphic entities, which is a very useful resource. Another interesting initiative was proposed by the Epigraphic Ontology Working Group (EpOnt[9]) and is focussed on building an application profile based on the concordance of ontologies, for recording epigraphic editions.

The analysis of all these models resulted in the definition of a preliminary shared set of entities related to inscriptions, graffiti and rock art that constitute the first draft of the future application profile, and

[3] http://www.cidoc-crm.org/crmarchaeo/
[4] http://www.cidoc-crm.org/crmtex/
[5] http://romanopendata.eu/sparql/doc/index.html
[6] https://www.eagle-network.eu/
[7] https://www.eagle-network.eu/wp-content/uploads/2013/06/EAGLE_D3.1_EAGLE-metadata-model-specification_v1.1.pdf
[8] https://www.eagle-network.eu/resources/vocabularies/
[9] https://groups.google.com/forum/#!forum/epont

will be validated in the following period when the collection level aggregation of information from this domain will be completed.

Scientific Data

WP4 has analysed several models for the definition of the application profile for scientific data. The scientific model developed by PIN and INFN seem very promising. The system is built around a general meta-model, flexible enough to provide descriptions, in a formal language, of the datasets produced by scientific research. Resulting metadata can then be re-encoded and published in multiple formats. The underlying metadata schema is inspired by CIDOC CRM principles for data modelling and maintains full compatibility with CIDOC CRM ontology to capture provenance and foster interoperability with Cultural Heritage information. The model is quite new and applicability tests are currently underway at different institutions to test its usability. The developers have also created a set of user interfaces to simplify and speed up the process of data gathering and metadata definition.

Currently, the full compatibility of this scientific model with AO-Cat has been verified, and many partners have confirmed its validity as a valuable application profile candidate. In the next period, the model will be tested during the scientific data aggregation phase.

4. Mapping to the ARIADNEplus Data Model

4.1 The X3ML framework

The X3ML framework is a set of small, open source, microservices that follow the SYNERGY Reference Model[10] of data provision and aggregation. They are designed with open interfaces and can be easily customised and adapted to complex environments. The key components of the X3ML framework to assist experts in completing the time-consuming and error-prone task of data provision and aggregation are:

- Mapping Memory Manager - 3M is a tool for managing mapping definition files. It provides a number of administrative actions to assist domain experts in managing their mapping definition files.
- 3M Editor – is an interface tool that allows domain experts to create mappings. It provides:
  - Source and target agnostic mapping facility
  - Guided mapping according to the logic of the deployed ontology
  - Comment and justification facility
  - Mapping storage
  - Separated instance generation practice for IT professionals
- X3ML Engine - The X3ML Engine realises the transformation of the source records to the target format. The engine takes as input the source data (currently in the form of an XML document), the description of the mappings in the X3ML mapping definition file and the URI generation policy file and transforms the source document into valid RDF, which corresponds with the input XML file, with respect to the given mappings and policy.
- RDF visualiser – The visualiser allows rapid inspection of the transformed records.

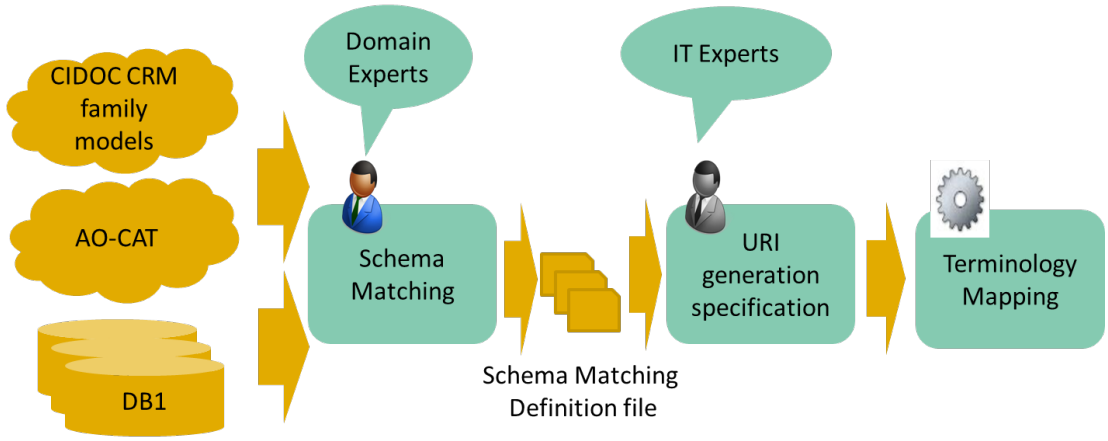


Figure 5. Data aggregation workflow.

In Figure 5, the data aggregation workflow is presented. Each ARIADNEplus data provider (domain expert) can use 3M to define the mapping of their native schema to the AO-Cat ontology. The first

[10] http://www.cidoc-crm.org/Resources/the-synergy-reference-model-of-data-provision-and-aggregation

step of the aggregation process is to define the schema matching, documented in a *schema matching definition file* which is both human and machine readable. The Mapping Memory Manager (3M) is a system to guide users to create appropriate matching and mapping files. The domain expert must identify the parts of the source schema to be used for the mapping to AO-Cat. It is not necessary, for the needs of the ARIADNEplus catalogue, to map to the entire schema, and only the parts of the schema that correspond to the AO-Cat are required. Data that may be confidential, or for internal use are not mapped. First, the domain expert identifies information of interest, and then the IT expert finds the corresponding tables/files to be aggregated. The 3M Editor provides a user friendly interface to create the matching definition file. The Source Schema Validator and the AO-Cat Schema Validator components assist the users in selecting valid paths with respect to the corresponding schemata, preventing mistakes.

The next step is the specification of the instance generation rules that define the URI generation specification for each AO-Cat class. This task is accomplished by the Instance Generator Rule Builder, which complements the Schema Matching Definition with the instance generation policies (based on the defined ARIADNEplus Generator Policy functions), producing the mapping definition (which is also called X3ML file). Both the schema matching definition and the mapping definition may be viewed with the Mapping Viewer. The files are stored to the Mapping Memory, which is an XML database. A set of generator policy functions have been defined that create appropriate URIs for all the classes of AO-Cat. For instance, the function that creates the URI for an AO_Collection is:

```
<generator name="CollectionURI-2" prefix="ariadneplus" shorten="yes">
    <pattern>Collection/{provider}/{term}</pattern>
</generator>
```

where {provider} is an argument that will take the name of the provider as a value, while {term} will take a valid identifier of the collection of interest as a value.

The ariadneplus namespace is defined as https://ariadne-infrastructure.eu/aocat/

The mapping definition and a sample input from the provider are given as input to the X3ML Engine component, which is responsible for transforming the records into the AO-Cat format. The transformation is tested online, within the environment of the 3M editor, and RDFvisualizer can be used to view the transformed records in real time. This is an easy way to verify and validate the mapping definition and to correct possible errors. When the user feels confident the mapping definition is complete, the full aggregation process can be initiated. The final result of the aggregation process is the full set of the provider's data in AO-Cat representation, so that they can be ingested into the ARIADNEplus cloud.

During the whole aggregation process domain experts are assisted by IT and AO-Cat experts. The steps are clearly defined and currently monitored by an online spreadsheet, but a dedicated Activity Dash is under development in the context of WP14, and will be used in the near future.

The X3ML Toolkit is deployed for ARIADNEplus (see Figure 6) on D4Science at:

https://ariadne.d4science.org/group/ariadneplus_mappings/mapping-tool

All users registered in the D4Science ARIADNEplus Mappings VRE can access it using their D4Science credentials.

Essential resources needed to start the aggregation workflow were produced and provided to all partners. These resources are continuously revised and updated:

Resource	Version as of 15 May 2020	Description	Access
AO-CAT	1.1.5	RDFS encoding of AO-CAT	https://data.d4science.net/w6Yw
ARIADNEplus GeneratorPolicy	1.4	Definition of URI and label generator functions	https://data.d4science.net/n7Zf
ARIADNEplus Template Mapping		Template mapping that loads AO-CAT 1.1.5, CIDOC CRM 6.2.1, CRMpe 3.1.2 and ARIADNEplus Generator Policy File 1.4 Can be used as a starting point by all partners.	https://ariadne.d4science.org/group/ariadneplus_mappings/mapping-tool Mapping 594
ARIADNEplus FastCat Template		Generic mapping from the FastCat export xml to AO-Cat. To be used by all partners that will create sample records via FastCat	https://ariadne.d4science.org/group/ariadneplus_mappings/mapping-tool Mapping 622
Manual	1.6	Section 5: Using the 3M Mapping tool	https://data.d4science.net/4HQr (See also *Bardi et al., 2020*)
X3ML Generators Manual	1.0	Mapping Memory Manager (3M) Instance and Label Generator Rule Builder	https://mapping.d4science.org/3M/Manuals/en/X3ML_Generators_Manual.pdf

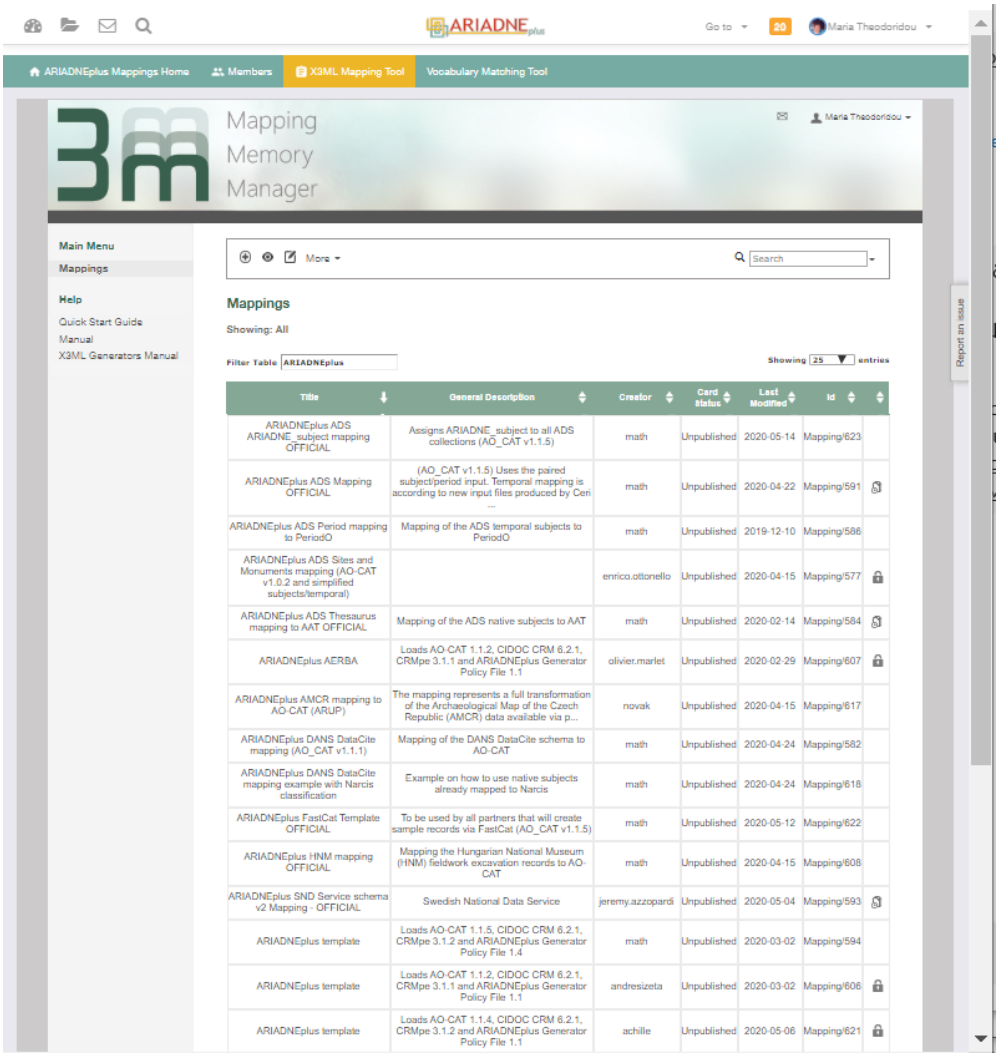


Figure 6. The X3ML framework deployment for ARIADNEplus.

A list of mappings to date implemented 3M is provided in section 7.4.

The 3M tool is used in combination with other data cleansing tools, developed by ARIADNEplus for the standardisation of vocabularies and the normalisation of temporal periods and time spans, to enrich and normalise archaeological data. These tools are presented in detail in section 4 of deliverable D5.2.

4.2 The FastCat tool

FastCat is an innovative tool designed for archaeologists, historians and other CH researchers who need to manually digitise structured and semi structured historical documents in a fast and accurate way, in order to create their research dataset. It combines ease of use and quick data entry of the classic spreadsheet functionality, with the information accuracy typically associated with a complex

database. This is accomplished through data entry templates, which are designed to mirror the structure and data entry logic of the original data source, in the digital space. A feature that makes FastCat stand out from traditional spreadsheet tools is that it can natively hold data tables within other data tables, which is very useful for recording historical data sources.

An ARIADNEplus resource catalogue template has been defined and implemented following the AO-Cat ontology.

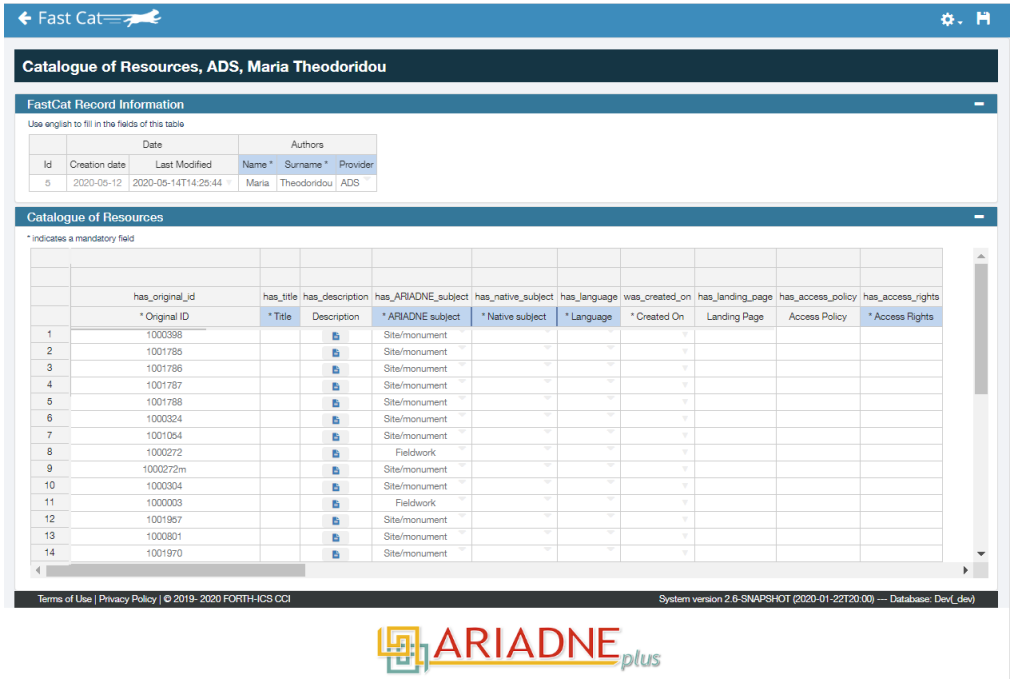


Figure 7. FastCat used to assign ARIADNE_subject to all ADS collections.

FastCat for ARIADNEplus is an easy way to familiarise data providers with the classes of AO-Cat and their use. Each data provider can map a few sample records from their collections to the AO-Cat and see how they are presented in the ARIADNEplus cloud.

5. Helpdesk for supporting data integration

As ARIANDEplus is a collaborative initiative with multiple participants representing varying levels of experience, it was important to assist content providers in a more structured way in all phases of their data contribution. Based on earlier experiences from the ARIADNE project, as well as other projects, it was decided that a ticketing service was needed where it is possible to offer efficient interaction with specialist teams of experts, and where one could track/document and follow up issues reported by participating members. In the D4Science infrastructure it is possible to implement the issue tracking system called “Redmine”. Redmine was selected because it integrates perfectly into the D4Science infrastructure, and provides seamless and transparent access to the Redmine help desk project with no need for further login.

Redmine is a web-based project management and issue tracking tool. Within Redmine it is possible to create, update, track, and resolve reported issues, while providing assistance at every stage of the process, from preparation to the definition of mappings, up to fine-tuning the data harvesting and data acquisition mechanisms within the ARIADNEplus infrastructure. In addition to the tracking system, it supports flexible role-based access control (to support the roles of the different users); Gantt charts provide visual representations of the project deadlines, linked to calendar; e-mail notifications and a wiki.

The service is up and running, allowing content providers to contact the right expert group, accelerating the process to find solutions. Additional benefits include fostering collaborations, and working as a future knowledge base with solutions to common problems and other information.

The helpdesk is primarily to be used if participants have queries to be resolved, features to request, suggestions, and bugs to report. When entering the Help Desk area there is an overview with information on how to use the system, for example the mandatory choice between Bug, Feature, and Support in the tracker field. Users must add information in the Subject and Description fields. The Category field is also mandatory, and allows users to choose between several areas described below. Choosing a category ensures the issue reaches the correct specialist team of experts.

Currently, the different categories are as follows:

WGS84: If the users have problems transferring coordinates to WGS84 and need advice/help, they are asked to describe the coordinate system they are using and the problem they are encountering when trying to transfer the coordinates.

Categorisation of your data: If users have questions about how to categorise their data and the information provided by WP4 (subtasks) is not adequate for their situation, they are asked to describe what type of data they have, and if possible, how they would like to categorise it.

Data supply – HowTo upload data: If the user needs help/advise on how to upload their data/metadata to the ARIADNE data cloud, they are asked to describe the type of data they have, the amount of data, if there is metadata that is required that they do not have, and any other relevant information. If they encounter problems when trying to upload their data/metadata, they are also asked to describe the type of problem.

FAST-CAT: If the users have questions regarding Fast-Cat, they are asked to try to specify the issue, and the help they require.

General questions: If the users have questions not related to any of the other categories, this is the category they can use. The user is asked to relate their question, then the helpdesk either will try to help them directly, or forward it to whoever might be able to solve the problem.

Mapping – CIDOC CRM: If the users have CIDOC CRM mapping issues, they are asked to specify the nature of the issue.

Mapping – PeriodO: If the users have issues related to PeriodO, they are asked to describe the problem and what type of advice/help they might need.

Mapping – Vocabularies: If the users have questions regarding mapping to the AAT and/or other vocabularies, they are asked to describe the problem, and what type of vocabulary they are using.

Mapping – 3M tool: If the user has issues with the 3M tool, they are asked to specify the issue is and the help they require.

Mapping – AO Cat Ontology: If users have issues with mapping AO Cat Ontology, they are asked to specify the issue, and what help they require.

Having a ticket system like Redmine will help keep track of progress until a problem is solved. It speeds up the process by automating some of the administrative tasks and giving the specialist teams more time to address and solve problems, making sure the community gets the help it needs when as quickly as possible. By creating ticket categories and assigning specialist teams of experts who are well versed in a problem, each category the knowledge from the experts can addressed more quickly. The ticket system will eventually contain enough information to populate a wiki, which when accessible to the community and in combination with other services provided by ARIADNEplus, will foster further spread of knowledge, common standards, and cooperation across international borders and designated communities. The creation of a helpdesk will assist the partners, and in future other data providers, in the integration work and in accessing and using the ARIADNEplus cloud. It will operate online and rely on the results of other Tasks of this WP and on the Portal setup.

6. Current status of the integration

This section reports on progress made on the integration of data and metadata provided by the ARIADNEplus archaeological data suppliers in the first 18 months of the project. Datasets follow an aggregation pipeline (described in D5.2) and are mapped to an underpinning ontology, the AO-Cat, using the 3M mapping tool (described in the previous section). Here we focus on discussing current and upcoming priorities, and reporting progress.

6.1 Aggregation priorities

Archaeology is a very broad domain, ranging from palao-anthropology to the study of standing buildings, and using techniques ranging from ancient DNA to LiDAR. ARIADNEplus seeks to integrate an extremely broad variety of archaeological data types. In the original description of work (T4.4, led by UoY-ADS) we defined these distinct sub-domains as follows (Sub-task leaders are named after each sub-task):

4.4.1 Paleo-anthropology (**CENIEH**)
4.4.2 Bio-archaeology and Ancient DNA (**FORTH (IMBB)**
4.4.3 Environmental Archaeology (**SND (SEAD)**
4.4.4 Inorganic Materials Study (**INFN**)
4.4.5 Dating (**INFN**)
4.4.6 Field Survey (**RUG**)
4.4.7 Archaeological finds made by general public (**AU**)
4.4.8 Remote Sensing (**ZRC-SAZU**)
4.4.9 Standing Structures **(LNEC**)
4.4.10 Spatio-temporal data (**ARUP-CAS**)
4.4.11 Maritime and underwater archaeology (**DGPC**)
4.4.12 Archaeological fieldwork (**INRAP)**
4.4.13 Inscriptions (**UB**)

In addition, we a core underpinning category has been added, inherited from ARIADNE, where it formed the bulk of the aggregated data:

4.4.0 Archaeological sites and activities **(UoY-ADS)**

This is now broken down into three sub-categories:

4.4.0.1 Site/monument
4.4.0.2 Fieldwork
4.4.0.3 Fieldwork report

To reflect data that partners wished to make available for integration we have also added:

4.4.14 Burials **(OEAW)**

Amongst the partners we have expertise covering most aspects, but existing vocabulary standards need to be identified, and a number of specialist application profiles and mappings to the A0-Cat (and thereby to the CIDOC-CRM) need to be developed. This is intensive, time consuming work, so in the first phase of the project it was essential to prioritise the data types and sub-domains to be tackled.

Prioritisation was guided by five criteria:

1. Ease of mapping and capacity to build upon lessons learned in ARIADNE, given experience with the previous data model, the ACDM.
2. Potential to aggregate a large number of records at an early stage. Analysis of datasets which partners were able to supply indicated that over two million records would be available for aggregation. This gives the portal developers in WP12 a critical mass of high quality data to use as test data in portal development.
3. The fact that these datasets correspond well with the “What”, “When”, “Where” parameters which have steered the development of the portal interface, and can be made interoperable by mappings to the Getty AAT (for “What”), use of Period (for “When”), and adherence to WGS 84 (for “Where”).
4. The fact that most partners had some datasets which fell into one or more of these parameters, and would therefore gain experience of the mapping process, and have some data visible in the portal by the end of the second year of the project.
5. User needs. Priorities were guided by the User Needs Survey conducted by SRFG in the first nine months of the project (D2.1).

The User Needs Survey compared the online accessibility of the data types against the perceived need for the ARIADNEplus portal to assist users to discover and access datasets or collections of the data types, resulting in the matrix shown in Figure 8.

		Discovery & access via the ARIADNEplus portal: Collection-Level (C-L) Rating of the helpfulness of support by the ARIADNEplus portal to discover and access datasets or collections of the data types Rating: % of Very helpful + Helpful (N = 341-355)		
		High	**Medium**	**Low**
Online Accessibility (OA) Rating of the online accessibility of the data types Rating: % of Very good + Good (N = 336-367)	**High**	Sites and monuments databases or inventories [OA:59.4%/C-L:93.2%] National GIS data & maps [OA:65.4%/C-L:93%]	Satellite or airborne remote sensing data (e.g. LiDAR) [OA:53.8%/C-L:89.2%]	
	Medium	Excavation data (e.g. excavation archive) [OA:47.8%/C-L:91.5%] Artefact/finds databases or image collections [OA:48.3%/C-L:90.4%]	Unpublished fieldwork reports ('grey literature') [OA:39.3%/C-L:88.1%] Field survey/ prospection data [OA:43.2%/C-L:87.7%]	Inscriptions, coins or other special databases [OA:41.7%/C-L:75.2%]
	Low	.	Environmental archaeology datasets [OA:36.2%/C-L:87.6%] Radiocarbon, dendrochronology and other dating data [OA:32%/C-L:84%]	Scientific data/analysis of biological remains [OA:29.2%/C-L:76.1%] Scientific data/analysis of inorganic remains [OA:28.1%/C-L:76.5%] Maritime and underwater archaeology data [OA:22.3%/C-L:57.8%]

Figure 8. Dataset and data collection discovery matrix.

The assessment of the five criteria led to the prioritisation of:

1. 4.4.0 The broad category of archaeological sites, monuments and activities (with its sub-tasks 4.4.0.1, 4.4.0.2 and 4.4.0.3).

Followed by upcoming priorities of:

2. 4.4.7 Archaeological finds: drawing on the growing European investment in national databases of finds made by members of the public across northern Europe and Scandinavia. These finds databases are rarely cross-searchable with sites and monuments databases, which for research users would provide an important infrastructure.
3. 4.4.14 Burials: given this also provides an early opportunity to aggregate data at item level, and these have a clear relationship with 4.4.0 and 4.4.7, but which can rarely be interrogated.

The interoperability of these three categories of dataset by the end of Year Two will provide a significant demonstration of the value of the ARIADNEplus infrastructure, allowing scholars to address, for example, major research questions of population migration in a spatial and temporal context.

6.2 First priority aggregation status

In line with the agreed priorities, we have spent time in the first 18 months of the project focussing on test data for Sub-task 4.4.0. Within Sub-task 4.4.0 each partner was ranked from 1-4, according to state of preparedness, as assessed at the two aggregation workshops held in Pisa, and subsequently proceeded according to that ranking, with variations according to staff availability.

The initial focus was on collections drawn from UoY-ADS, which holds many datasets that fall within the 4.4.0 category, but with a few additional sub-types. This has allowed us to test the aggregation pipeline, amending as necessary and to refine the AO-Cat according to the data provided. It has also allowed the core team to become familiar with the key issues to be encountered during the project. The learning curve means there is now a stable process which should make aggregation easier for less experienced partners.

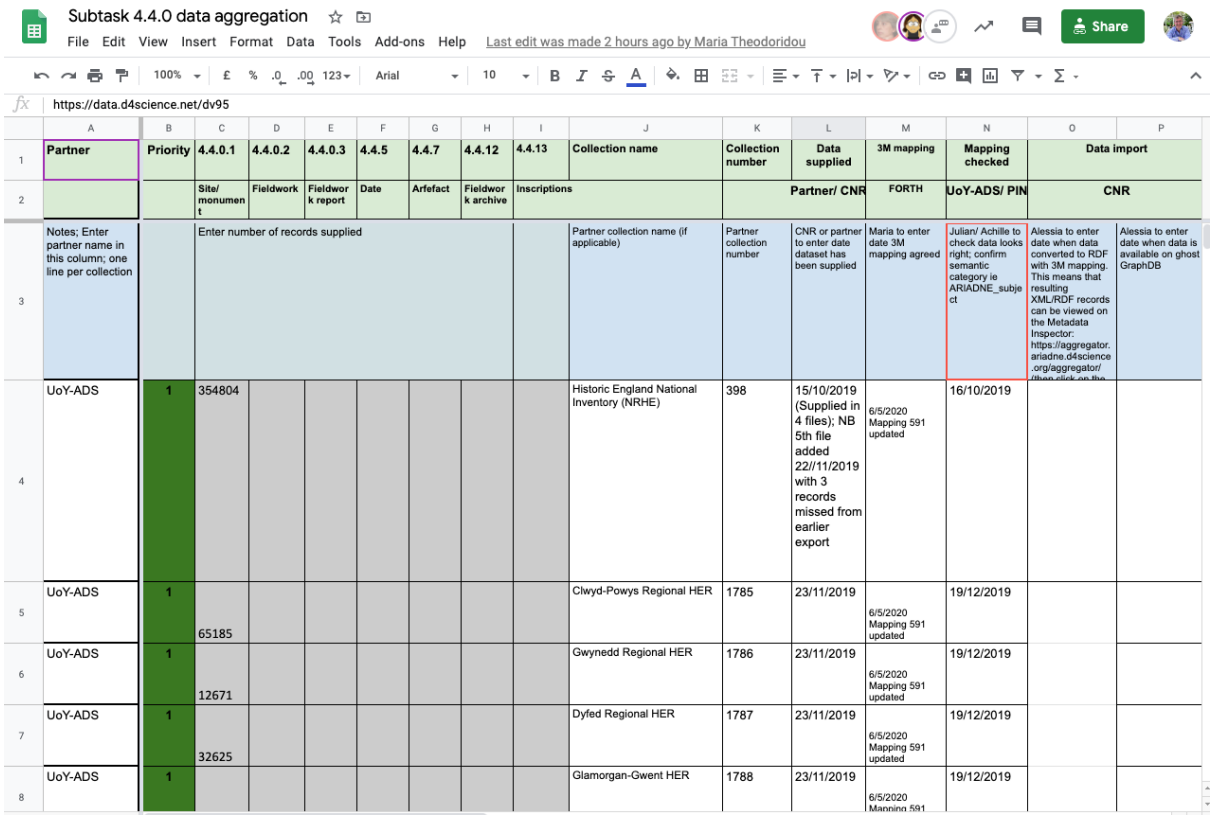
Table 1 shows all UoY-ADS collections, which by May 2020, had completed all stages of the aggregation pipeline, including 3M mapping, AAT mapping, cleaning of temporal data, and transformation into RDF. This comprises c. 1.3 million records in total.

Having successfully aggregated the ADS data we have now moved onto datasets from other partners within Groups 1 and 2 (Table 2).

Table 1: UoY-ADS collections aggregated as of May 2020

Sub-type	No of records	Collection name	Collection number
4.4.0.1	354804	Historic England National Inventory (NRHE)	398
4.4.0.1	65185	Clwyd-Powys Regional HER	1785
4.4.0.1	12671	Gwynedd Regional HER	1786
4.4.0.1	32625	Dyfed Regional HER	1787
4.4.0.1	23994	Glamorgan-Gwent HER	1788
4.4.0.1	20173	Defence of Britain Archive	324
4.4.0.1	971	Exeter City HER	1054
4.4.0.2	14152	Greater London Sites and Monuments Record	272_events
4.4.0.1	54805	Greater London Sites and Monuments Record	272_monuments
4.4.0.2	7119	Historic England NRHE Index to Microfilmed Archaeological Archives	3
4.4.0.1	1724	Isle of Man HER	1957
4.4.0.1	15437	Lake District National Park HER	801
4.4.0.1	16205	Merseyside HER	1970
4.4.0.1	3169	Museum of London Archaeological Archive	270
4.4.0.1	328747	National Monuments Record of Scotland	1
4.4.0.7	4610	National Museums and Galleries of Wales Collection	397
4.4.0.1	63197	National Trust HBSMR	328
4.4.0.1	16463	Northern Ireland Sites and Monuments Record	276
4.4.0.1	13497	Northumberland Sites and Monuments Record	292
4.4.0.1	6997	Parks and Gardens UK	858
4.4.0.1	100	Shetland Sites and Monuments Record	273
4.4.0.1	29137	Somerset Historic Environment Record	388
4.4.0.1	6997	South Gloucestershire SMR	321
4.4.0.2	4092	The Geophysical Survey Database	1091
4.4.0.1	9037	Warwickshire Sites and Monuments Record	367
4.4.0.1	32856	West of Scotland Sites and Monuments Record	269
4.4.0.2	998	York Archive Gazetteer	271
4.4.0.2	6830	Yorkshire Dales National Park Historic Environment Record	420_event
4.4.0.1	23961	Yorkshire Dales National Park Historic Environment Record	420_monument
4.4.0.3	56328	GLL	1093
4.4.0.5	8993	Archaeological Site Index to Radiocarbon Dates from Great Britain and Ireland	4
4.4.0.13	1578	England's Rock Art	836
4.4.0.13	20188	Historic Milestones database	1972
TOTAL	1257640		

The status of aggregation is currently being managed through shared access by the aggregation team to a Googlesheet dashboard, in which the progress through the pipeline can be monitored (Figure 9). In due course this will be migrated to the Activity Dash online tool, being developed by FORTH.



Partner	Priority	4.4.0.1	4.4.0.2	4.4.0.3	4.4.5	4.4.7	4.4.12	4.4.13	Collection name	Collection number	Data supplied	3M mapping	Mapping checked	Data import	
		Site/ monument	Fieldwork	Fieldwork report	Date	Arfefact	Fieldwork archive	Inscriptions			Partner/ CNR	FORTH	UoY-ADS/ PIN	CNR	
Notes; Enter partner name in this column; one line per collection		Enter number of records supplied							Partner collection name (if applicable)	Partner collection number	CNR or partner to enter date dataset has been supplied	Maria to enter date 3M mapping agreed	Julian/ Achille to check data looks right; confirm semantic category ie ARIADNE_subject	Alessia to enter date when data converted to RDF with 3M mapping. This means that resulting XML/RDF records can be viewed on the Metadata Inspector: https://aggregator.ariadne.d4science.org/aggregator/	Alessia to enter date when data is available on ghost GraphDB
UoY-ADS	1	354804							Historic England National Inventory (NRHE)	398	15/10/2019 (Supplied in 4 files); NB 5th file added 22//11/2019 with 3 records missed from earlier export	6/5/2020 Mapping 591 updated	16/10/2019		
UoY-ADS	1	65185							Clwyd-Powys Regional HER	1785	23/11/2019	6/5/2020 Mapping 591 updated	19/12/2019		
UoY-ADS	1	12671							Gwynedd Regional HER	1786	23/11/2019	6/5/2020 Mapping 591 updated	19/12/2019		
UoY-ADS	1	32625							Dyfed Regional HER	1787	23/11/2019	6/5/2020 Mapping 591 updated	19/12/2019		
UoY-ADS	1								Glamorgan-Gwent HER	1788	23/11/2019	6/5/2020 Mapping 591	19/12/2019		

Figure 9: The Googlesheet dashboard used in Task 4.4.

Table 2: Collections of Sub-type 4.4.0 aggregated from other partners

Partner Name	Collection	No of records
HNM	Hungarian National Museum Archaeology database	16257
AIAC	Fasti Online	c.4200
DANS-KNAW	eDNA	c.51000
SND	Swedish National Data Service	c.800
ARUP	AMCR	c.100000

Several other partners with small numbers of records have begun to use the Fast-Cat tool for the manual entry of up to 20 records, and this has been adopted by CONICET, INFN and FORTH (IMBB).

6.3 Forthcoming priorities

From June, until the end of 2020, we plan to complete the initial aggregation of data within the 4.4.0 category (Table 3), as well as moving onto the other sub-tasks identified as priorities (i.e. 4.4.7 and 4.4.14) (Table 4). Partners within these groups have already started developing mappings to AO-Cat, cleaning data and standardising vocabularies.

Table 3 Next 4.4.0 priorities

Partner Name	Collection	No of records
IAA	Atiqot - fieldwork reports	3490
FI	Isleif - Icelandic NMR	c.100000
ROCEEH (Associate Partner)	ROADS	
ASU	tDAR: https://core.tdar.org/search	c.20000
NARA	Nabunken: https://sitereports.nabunken.go.jp/en	c.58000
MIBAC-ICCU		c.36000
INRAP	Dolia	c. 43000
CARARE	Discovery programme and other sources, Ireland	c. 180,000

Table 4 Priorities from other sub-tasks

Partner Name	Sub-task	Collection
UoY-ADS	4.4.7	Portable Antiquities Scheme
DANS-KANW	4.4.7	PAN
AU	4.4.7	DIME
UH	4.4.7	Findsampo
RGK	4.4.7	Roman coins database
ZRC-SAZU	4.4.14	N/A
OEAW	4.4.14	N/A
ARUP	4.4.14	N/A

Finally, in late 2020, we will develop an update procedure and schedule for refreshing existing data sets, to ensure the ARIADNE infrastructure is kept up-to-date.

Beyond 2020, as we enter the second phase of the project, we will turn our focus to the other sub-tasks, where there may only be a single collection level record in the ARIADNE portal, but where item level integration will be achieved via our Linked Open Data Cloud and the VREs.

6.4 AO-Cat mappings

Each collection requires a mapping to the AO-Cat, generally using the 3M tool, although where Fast-Cat has been adopted there is a single Fast-Cat 3M mapping to streamline the process. For other partners we have tried to minimise the proliferation of mappings, to avoid redundancy and the creation of overheads for sustainability, because if the AO-Cat is modified, the mappings will also need to be updated. For example, at one stage there were over ten different mappings to cover the range of UoY-ADS datasets, which we have reduced to four.

As of May 15, 2020 the following mappings have been implemented in 3M:

Resource	Mapping
UoY-ADS	
Generic mapping covering all ADS collections	591
Mapping ADS native subjects to AAT	584
Mapping assigning ARIADNE_subjects to all ADS collections	623
Mapping of the ADS temporal subjects to PeriodO	586
HNM	
Hungarian National Museum Archaeology database	608
AIAC	
Fasti Online	614
DANS-KNAW	

DataCite	582
Example on how to use native subjects already mapped to Narcis	618
SND	
Swedish National Data Service	593
ARUP	
AMCR	617

7. Digital Libraries integration

In this chapter, we present the activities related to Task 4.5, "Integrating archaeological digital libraries with ARIADNEplus" which are aimed at linking the ARIADNEplus Data Infrastructure with repositories of scientific publications, exploiting, in particular, OpenAIRE and its open access digital archives and the links to individual journals such as Internet Archaeology or A&C. The task will also explore how the ARIADNEplus text mining service (Task 17.4) may be used to improve the metadata of the published papers. Note that T4.5 activities started at the end of M11 and thus the work of this task is still in its beginnings.

The OpenAIRE initiative[11] is a legal entity (OpenAIRE AMKE) representing a partnership of more than 50 institutions (in Europe and beyond) working to promote and support a sustainable implementation of Open Access and Open Science policies for reproducible science, transparent assessment and omni-comprehensive evaluation.

OpenAIRE features a technical infrastructure that features an open, de-duplicated, participatory metadata research graph of interlinked scientific products (including research literature, datasets, software, and other types of research products like workflows, protocols and methods), with access rights information, linked to funding information, research communities and infrastructures.

The resulting graph is called the OpenAIRE Research Graph[12]. Added-value services are also built on top of the graph to offer Open Science services to different stakeholders of the scholarly communication ecosystem, including researchers, content providers (e.g. repository, journals), research communities, research infrastructures, research organisations, and funders.

OpenAIRE has already set up a *gateway for Digital Humanities and Cultural Heritage* that can be found in beta version here: https://beta.dh-ch.openaire.eu (Figure 10). Its aim is to offer a single-entry point for discovering research products relevant in the domain of Digital Humanities and Cultural Heritage available in the OpenAIRE Research Graph. This broad definition includes Humanities, Cultural Heritage, History, Archaeology and related fields. Researchers will also find in the gateway Open Science publishing tools to (i) manually deposit research products of any type (i.e., not only scientific literature but also datasets, software, workflows, methods, etc.) and get a persistent identifier via Zenodo[13]; (ii) add details of the research context of a research product by creating links to projects and to other related research products (e.g. a link between a dataset and the software that analysed it); and (iii) manually add research products to the community gateway. Via such tools, researchers populate and access an open, participatory scholarly communication graph of interlinked objects through which they can share any kind of products in their community, maximise re-use and reproducibility of science, and outreach the scholarly communication at large. The graph is also analysed for the calculation of statistics that support the monitoring of the uptake of Open Science principles among its researchers.

[11] https://www.openaire.eu
[12] https://zenodo.org/communities/openaire-research-graph
[13] https://zenodo.org

The gateway is configurable by a set of "managers", who are experts in the domain and aware of the needs of the community. Managers can configure the algorithms used by the OpenAIRE infrastructure to assign scientific products in the OpenAIRE Research Graph to the community, based on information available in the metadata or in Open Access publications, including full-texts (full-text mining). OpenAIRE algorithms analyse the research graph to identify all products that are relevant for the community (e.g. products with relevant subjects, deposited in thematic or community-specific repositories and archives, linked to domain-specific projects). In addition, managers can also act as "community moderators" by accepting and rejecting links and products added manually to the gateway by researchers, and they can select which statistics and charts are relevant and should, therefore, be publicly visible in the gateway.

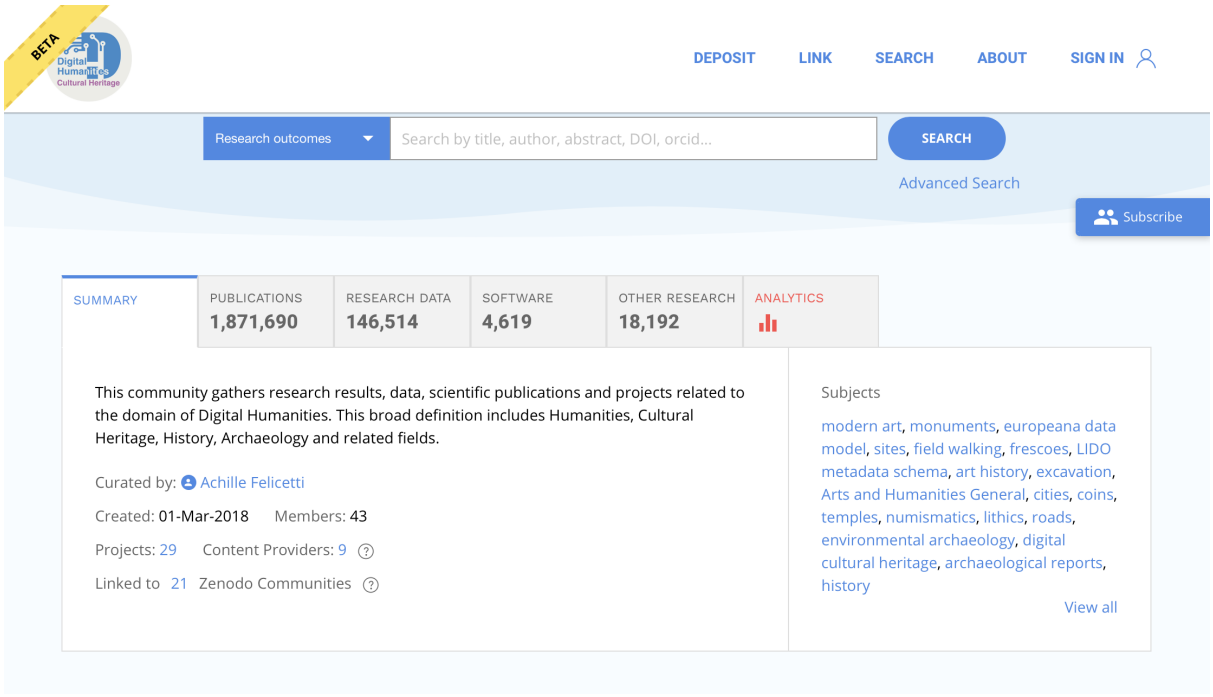


Figure 10: The Digital Humanities and Cultural Heritage community dashboard.

7.1 Gateway configuration

In this task, we will improve the initial configuration of the Digital Humanities and Cultural Heritage gateway for the needs of the ARIADNEplus community by updating the list of relevant projects, content providers, and relevant organisations. In addition, we plan to harvest metadata from ARIADNEplus to look for links to these datasets within publications, which can also be added to this gateway (see next section).

Currently, the list of relevant projects is 30, and they are funded by a number of institutions, including: 17 projects by the European Commission, five by Research Council UK (RCUK), three by the Australian Research Council (ARC), two by the Austrian Science Fund (FWF), two by the National Science Foundation (NSF), USA, and one by the Ministero dell'Istruzione dell'Università e della Ricerca, Italy (MIUR). The list can be seen here: https://beta.dh-ch.openaire.eu/search/find/projects, and of course includes the EC projects ARIADNE and ARIADNEplus (see Figure 11). As a result, all research outcomes

associated with these projects appear in the gateway, either because they have been linked/claimed by the projects, or because of associations created by OpenAIRE's full-text mining algorithms.

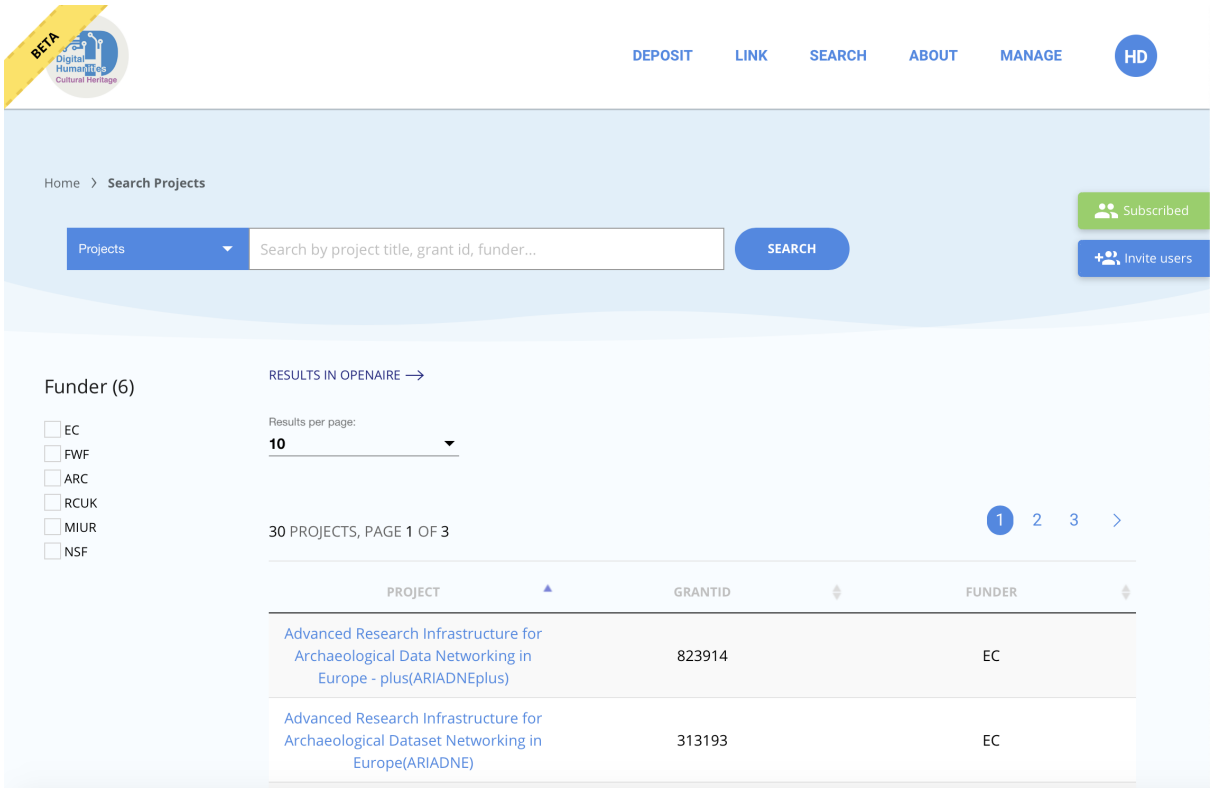


Figure 11: Relevant projects funded by the EC and other institutions, used in the configuration of the gateway.

The list of relevant projects will need to be updated for the purposes of ARIADNEplus. OpenAIRE has already integrated metadata from almost three million projects funded by 29 different funders from Europe and beyond. The full list of funders and projects can be viewed on the OpenAIRE Explore portal at the URL https://beta.explore.openaire.eu/search/find/projects. Thanks to the administrative panel of the gateway, managers will be able to search for relevant projects across all 29 funders by using specific keywords in the dedicated text box. The tool will return all the projects whose titles, acronyms (and soon abstracts) contain the user search terms. Figure 12 shows, for instance, the result of a search with the term "archaeology", which returns 52 EC (and 7069 projects from all funders in OpenAIRE) that are likely relevant in the domain of Archaeological Research and can thus be added to the configuration by clicking on the "plus" button on the right.

The list of relevant content providers currently includes nine repositories, as shown in Figure 13. OpenAIRE already integrates more than 12K content providers world-wide, aggregating their metadata records directly or via an aggregator (e.g. DOAJ and DataCite). Managers of the gateway can search for relevant providers and add them to the configuration at any time via the administrator web GUI. As shown in Figure 14, a search for the keyword "archaeology" returns 110 content providers that managers can evaluate for inclusion in the gateway configuration. For instance, the journals "Open Archaeology" and "Archaeology International" are probably very relevant to Archaeological Research, while "JIIA Eprints Repository" might be a relevant thematic repository to add.

Since the journal "Internet Archaeology" is already mentioned in the DoA of ARIADNEplus, we have recently added it to the list by using the "+" blue circular button on the right of the entry (Figure 14); it can be removed again by a manager by pressing the "x" red circular button.

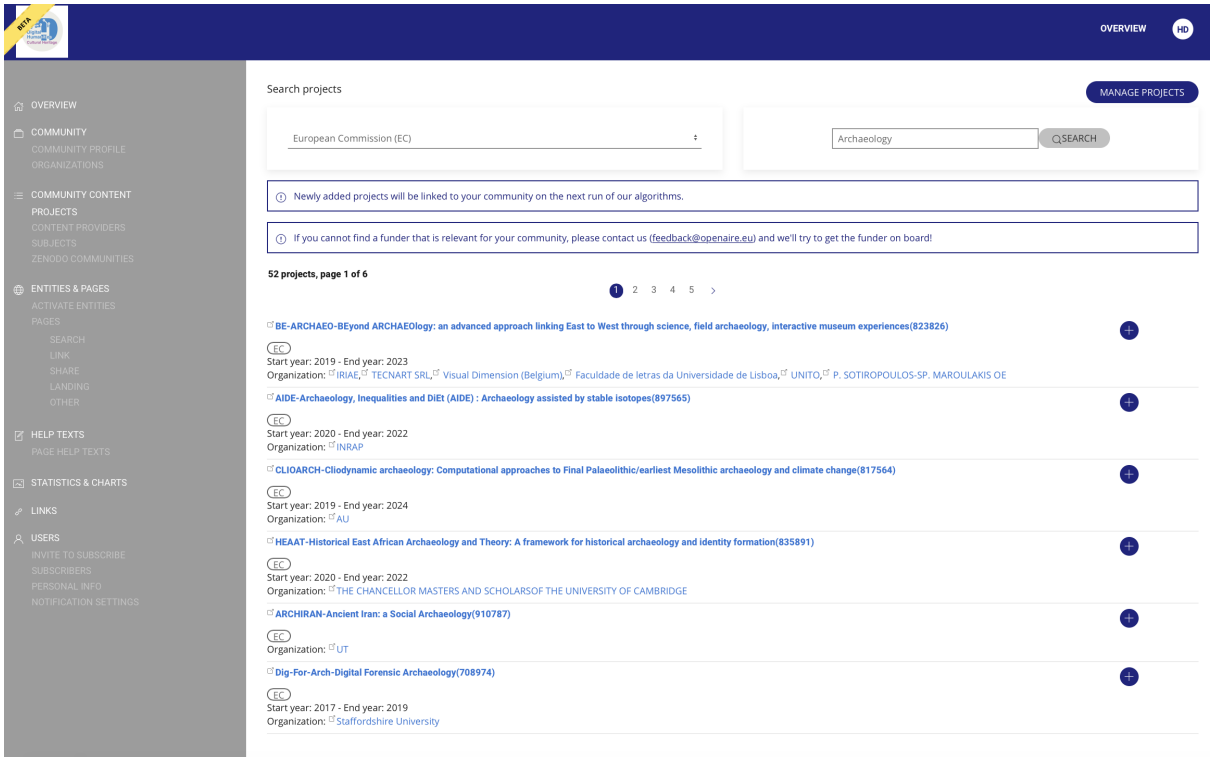


Figure 12: Improving the coverage of the research products in the gateway by adding more projects to the gateway configuration.

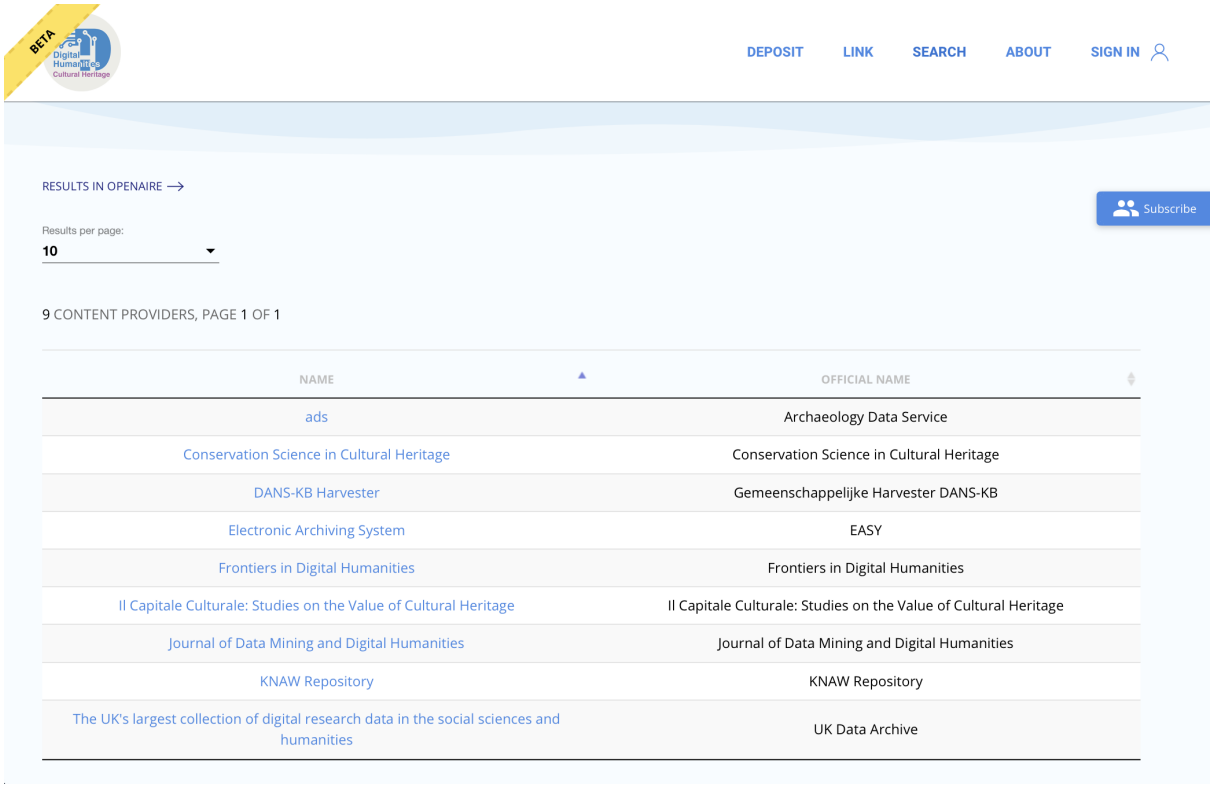


Figure 13: Current content providers used in the configuration of the gateway.

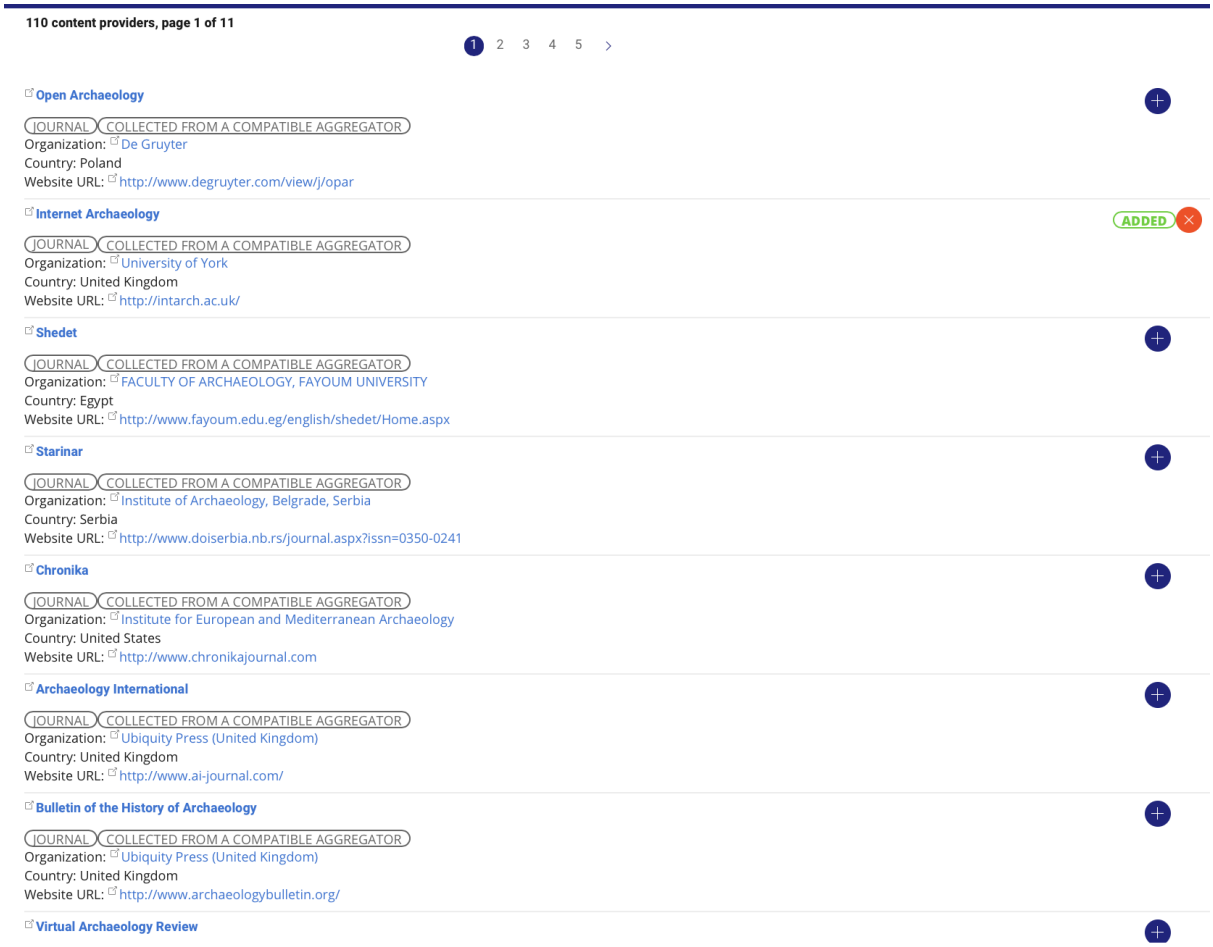


Figure 14: Improving coverage of research products in the gateway by finding relevant content providers among those aggregated by OpenAIRE.

If a content provider relevant to a community cannot be found, it probably is not OpenAIRE compliant. However, managers should feel free to contact OpenAIRE (feedback@openaire.eu) so that the situation can be investigated, and OpenAIRE may be able to get the provider to participate.

The gateway can also be configured to include research products whose keyword list (subjects) contains at least one specified in the configuration. See list of current subject configuration in Figure 15 below.

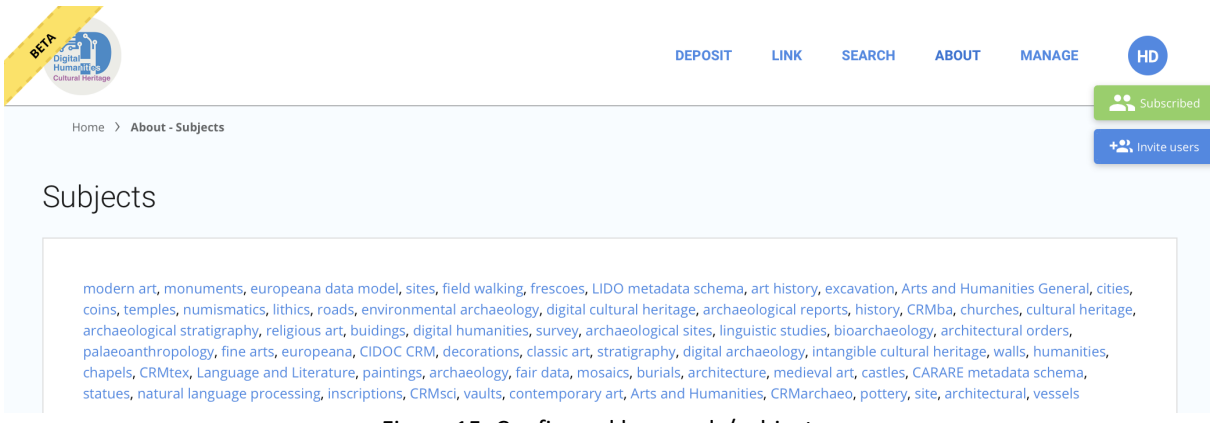


Figure 15: Configured keywords/subjects.

If one keyword from a research product in the OpenAIRE Research Graph is equal (ignoring case) to one of the terms in the list above, the product is included in the gateway. For example, 442 research outcomes are associated with the word “monuments” (271 publications, 155 research data such as images, and 16 other products), as shown in Figure 16. This approach is very basic and addresses obvious cases for disciplinary research communities. The effectiveness of the algorithm mainly depends on the established practices of researchers when depositing their products (i.e., if they use keywords and terms from a standard vocabulary). When such practices are not well established, a more intelligent approach must be adopted, such as domain-specific classifiers utilising advanced full-text mining techniques.

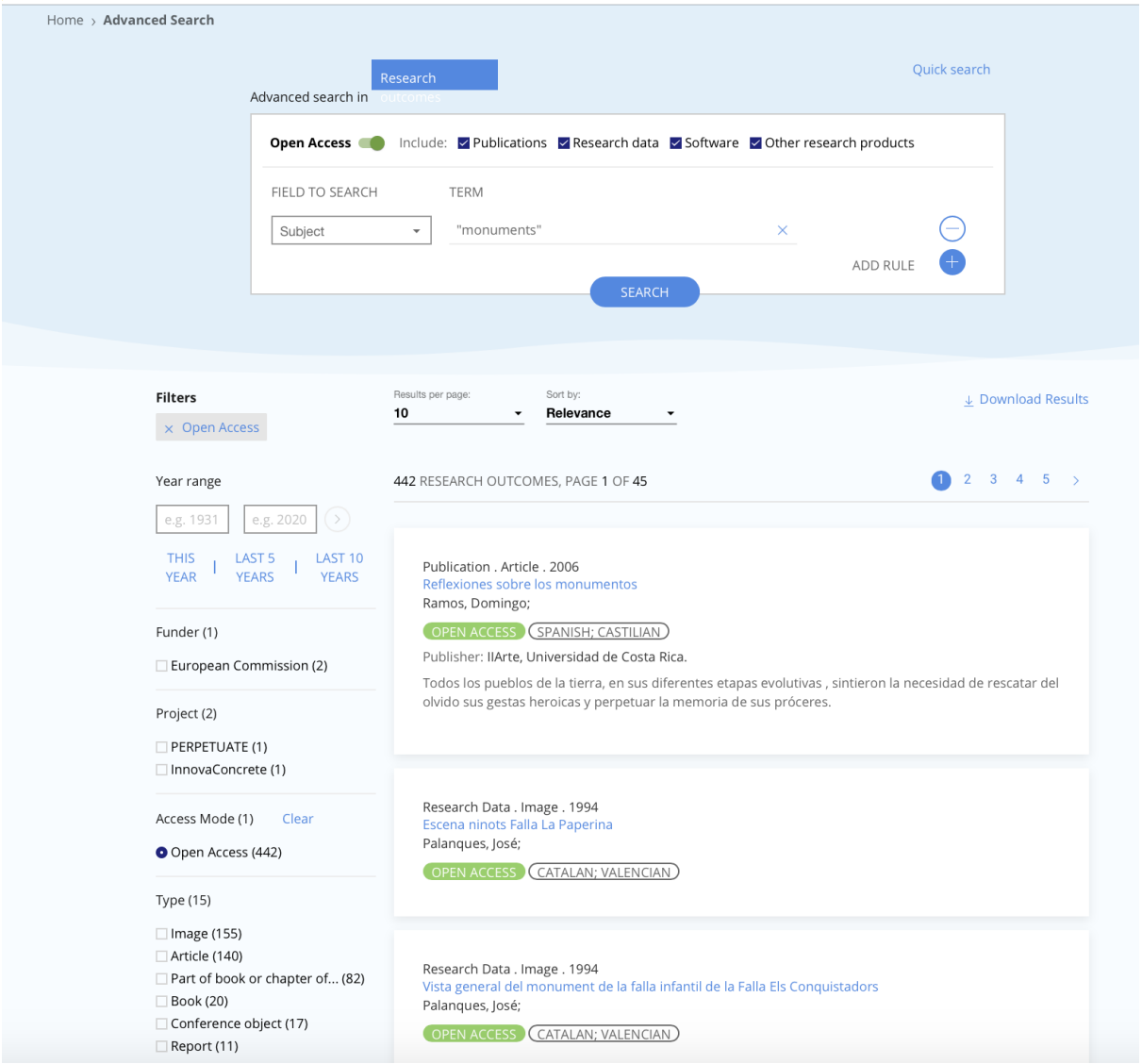


Figure 16: Research outcomes associated with the subject “monuments”.

Finally, the gateway can be configured for Zenodo communities. Zenodo is a multi-disciplinary repository for sharing, preserving and showcasing research products of any type (e.g. publications, posters, presentations, research data, software). When a user deposits a product on Zenodo, a DOI (if not already available) is assigned to it and the metadata are pushed in real-time into the OpenAIRE Research graph. In order to group depositions based on user-defined criteria, Zenodo introduces the

concept of “Zenodo communities”. They can be seen as “curated containers” of research products. Often, Zenodo communities are used by project managers to group deliverables and outputs of their research products, by institutions that do not have an official open access institutional repository, or by research groups working on a specific research topic.

Community gateway managers can search for Zenodo communities whose products are relevant to the domain and add them to the configuration: all research products deposited in the selected Zenodo communities will then be discoverable via the gateway. In addition, Zenodo communities selected by the managers are suggested to the users of the gateway in the “Deposition” section. Managers can also select or request to create one “main” Zenodo community to be associated with the gateway (e.g., as shown in Figure 17, where there is the “OpenAIRE-Connect Digital Humanities and Cultural Heritage” master Zenodo community). Typically, one of the gateway managers is also the curator of the main Zenodo community.

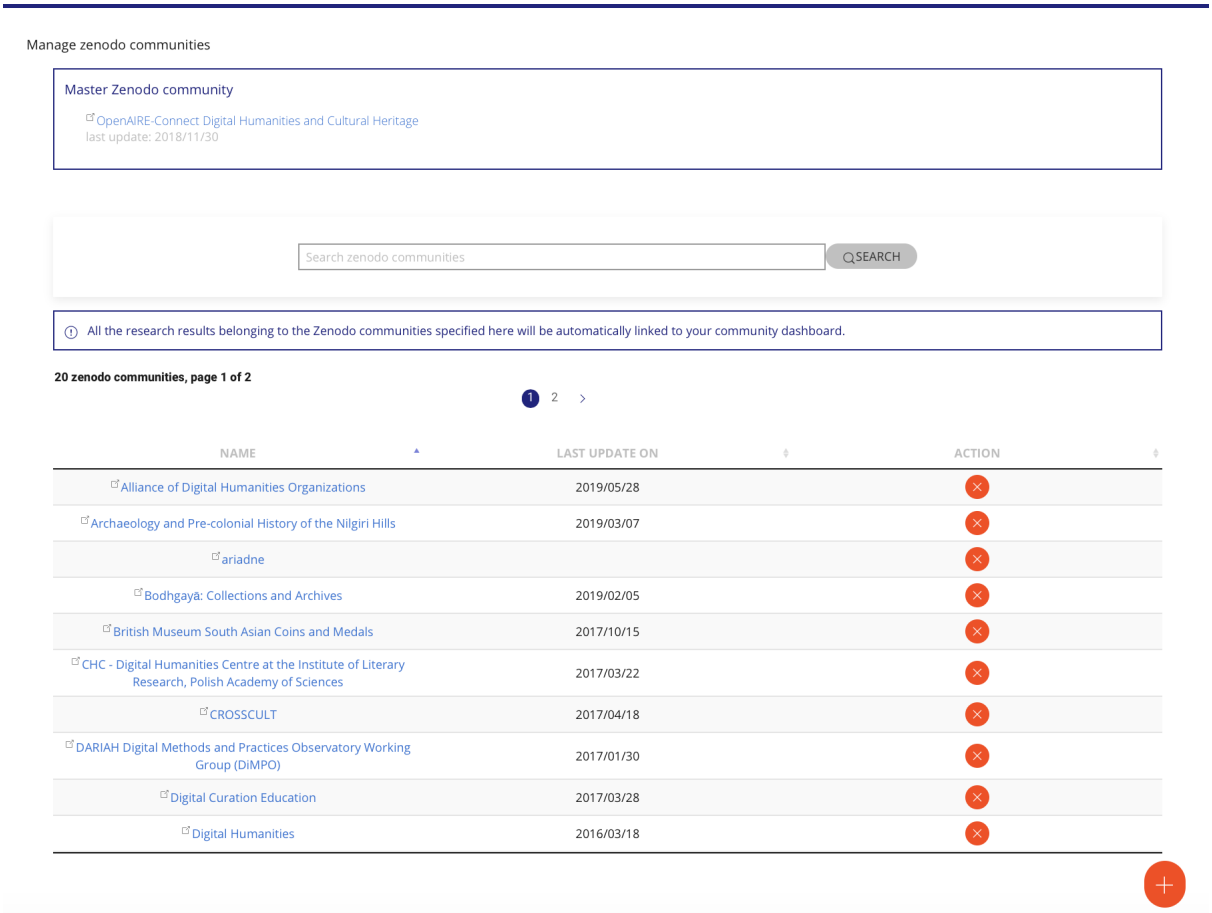


Figure 17: Managing Zenodo communities in the gateway.

With a quick search using the keyword “archaeology”, it is possible to get a list of 40 possibly relevant Zenodo communities (Figure 18). The list needs to be analysed by gateway managers so that only the communities that are relevant are added to the configuration. In addition, managers can try different keyword searches to identify additional interesting communities.

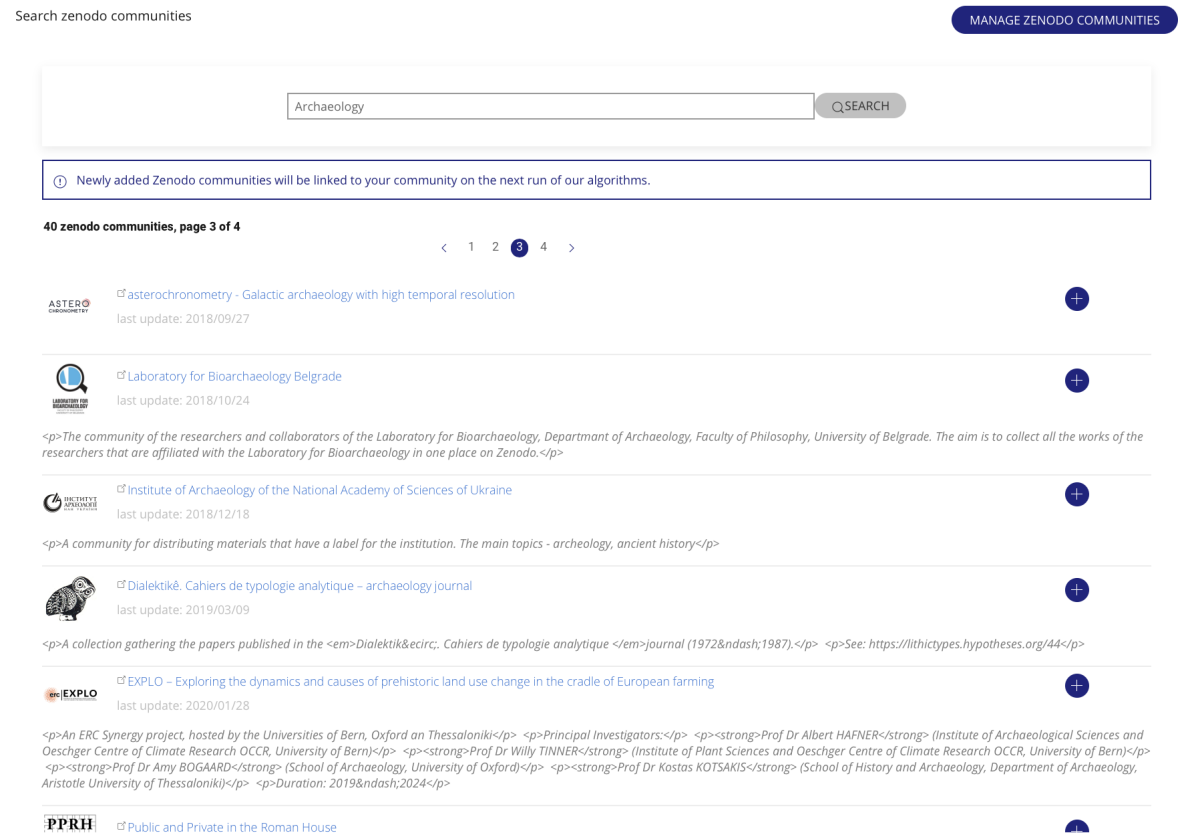


Figure 18: Improving the coverage of the research products in the gateway by identifying discipline-specific Zenodo communities.

7.2 Linking ARIADNEplus datasets to publications

As mentioned earlier, we also want to harvest dataset metadata from ARIADNEplus to look for links from OpenAIRE publications to these datasets. This will be a continuation of work started during the first ARIADNE project, and described in Section 3.1 of deliverable D16.3: http://legacy.ariadne-infrastructure.eu/resources-2/deliverables/d16-3-final-report-on-data-mining/.

The aim was to build a bridge between ARIADNE and OpenAIRE. In ARIADNE, we implemented an algorithm for citation extraction and algorithm dataset matching, as well as clustering and similarity algorithms, implemented on top of our madIS[14] system. Using these algorithms, we ran a number of experiments, but the most promising results were produced when we ran the citation extraction algorithm on ADS grey literature reports searching for hooks to OpenAIRE publications: 216 direct citation links from ADS reports to PubMed publications were found, as well as 83 indirect links. The direct links were all high confidence links, the indirect links were generally medium confidence; seven false positives were removed by the confidence threshold, which was optimised after manual curation. In total, 299 valid citation links from ADS grey literature reports to PubMed were found, mostly references to publications in medical history journals/epidemiology (e.g. yellow fever in the

[14] https://github.com/madgik/madis

1790s), geochronology, paleopathology, dental anthropology, anatomy, DNA studies, and so on. Indirect links were usually from an ADS report to a review of the cited work in OpenAIRE. So while not a direct citation match, still very relevant. In other words, an ADS report was linked to an OpenAIRE publication which discussed, presented or reviewed the publication cited by ADS, but is not the actual publication.

All links found were given for further assessment to an archaeologist, to confirm their relevance to ARIADNE. According to the archaeologist's summary evaluation report, "the vast majority of records are absolutely useful and relevant". The archaeologist also noted that topics were related to fields such as anthropology, biology, palaeontology, pottery, and also outside the UK, such as Oceania. With the exception of a few records, they were usable: 292 of the links were judged as relevant, six links as possibly relevant, and eight links as irrelevant. This was in line with our own internal evaluation.

Now, in ARIADNEplus, we plan to go much further than finding links from grey literature to OpenAIRE publications; we want to look at all ARIADNEplus datasets and objects and see if they can be linked to publications by investigating if publications cite any of the ARIADNEplus data. Such relationships can then be used to link content from OpenAIRE to the ARIADNEplus portal.

In addition, we want to explore how the ARIADNEplus text mining and NLP/NER (Named Entity Recognition) service (Task 17.4) may be used to improve the metadata of the published papers. The service will be based on the previous ARIADNE text mining tool, further developed into TEXTCROWD[15], a cloud-based NLP tool created as a Science Demonstrator within the EOSCpilot EU project. However, this collaboration between tasks T4.5 and T17.4 activity has not yet begun, as these are dependent upon T17.4 developments and the mid-term interim report on ARIADNEplus service - JRA4, deliverable D17.1 at month 24. Thus, we will describe our work progress and results relevant to this inter-task collaboration in deliverable D4.3.

[15] https://textcrowd.d4science.org

8. Conclusions

In conclusion, all the activities of WP4 in this first phase of the project contributed significantly to the integration of the archaeological datasets in ARIADNEplus. The AO-Cat model was the most remarkable result of these activities as its development benefited from the contribution of all partners and highlighted the excellent synergy between them. The definition of this standard made it possible to achieve an excellent level of integration already during the first 18 months of the project. This was favoured by the IT platform and the tools made available to implement the AO-Cat model and make it effective.

The mapping, enrichment and conversion tools specifically created for this purpose have played an important role in assisting content providers in all phases of the preparation and contribution of their datasets. The use of the Helpdesk also simplified and made more efficient the collaboration between technicians and content providers. Significant progress has also been made on the integration of the ARIADNEplus data space with repositories of scientific publications, to establish links between the archaeological data in ARIADNEplus and the information present in external digital libraries. The use of advanced text mining tools has made this activity particularly effective.

During the second phase of the project, WP4 will continue its standardisation and data integration activity in order to extend it to all content providers and complete this task. The AO-Cat model will be further extended and made compatible with the new application profiles already being defined within the various research domains. As a result, the ARIADNE Data Model will be released and used for extending the semantic integration also to the item level. This will foster compatibility and interoperability of the ARIADNEplus platform with other important European research infrastructures such as EOSC.

9. References

Aloia, N., Binding, C., Cuy, S., Doerr, M., Fanini, B., Felicetti, A., Fihn, J., Gavrilis, D., Geser, G., Hollander, H., Meghini, C., Niccolucci, F., Nurra, F., Papatheodorou, C., Richards, J., Ronzino, P., Scopigno, R., Theodoridou, M., Tudhope, D., Vlachidis, A. and Wright, H., "Enabling European Archaeological Research: The ARIADNE E-Infrastructure", *Internet Archaeology* 43 2017, https://doi.org/10.11141/ia.43.11

Aloia, N., Debole, F., Felicetti, A., Galluccio, I., Theodoridou, M.,, "Mapping the ARIADNE Catalogue Data Model to CIDOC CRM: Bridging Resource Discovery and Item-Level Access", SCIRES-IT SCIentific RESearch and Information Tech-nology, Vol. 7, 2017, http://dx.doi.org/10.2423/i22394303v7n1p1

Felicetti, A., Niccolucci, F., "A CIDOC CRM-based Model for the Documentation of Heritage Sciences". Proceedings of the 3rd International Congress on Digital Heritage (DH2018). Institute of Electrical and Electronics Engineers, 2018, https://doi.org/10.1109/DigitalHeritage.2018.8810109

Felicetti, A., Williams, D., Galluccio, I.; Tudhope, D., Niccolucci, F., "NLP Tools for Knowledge Extraction from Italian Archaeological Free Text". Proceedings 3rd International Congress on Digital Heritage (DH2018). Institute of Electrical and Electronics Engineers, 2018, https://doi.org/10.1109/DigitalHeritage.2018.8810001

Marketakis, Y., Minadakis, N., Kondylakis, H., Konsolaki, K., Samaritakis, G., Theodoridou, M., Flouris, G., & Doerr, M., X3ML mapping framework for information integration in cultural heritage and beyond. Int J Digit Libr 18, 301–319 (2017). https://doi.org/10.1007/s00799-016-0179-1

Bardi, A., Binding, C., Felicetti, A., Meghini, C., Richards, J. & Theodoridou, M. Unpublished report 2020. Data Aggregation Pipeline: User Guide. Version 1.7. 24/4/2020.

Niccolucci, F. and Richards, J.D. 2019 "ARIADNE and ARIADNEplus" in Richards and Niccolucci (eds.) The ARIADNE Impact.Archaeolingua, 7-25

Appendix

AO-Cat

This Section presents the full AO-Cat ontology, with the complete set of axioms. Wherever possible, axioms are given in the functional-style syntax of OWL 2 DL. The axioms not expressible in OWL 2 DL are given in first-order logic.

Class axioms

Class axioms consist of:

- Subclass Axioms
- Equivalent Classes
- Disjoint Classes
- Disjoint Union of Class Expressions

These kinds of axioms will be considered in the rest of this Section, with the exception of the axioms concerning the cardinality of object properties. These axioms are expressed as subclass or equivalent class axioms on class expressions, so they should belong in this Section. However, for readability these axioms will be given in the next Section, along with the axioms that define the object properties whose cardinality they concern. In contrast, domain and range axioms, which belong to next Section, are also redundantly given in this Section, for convenience.

Subclass and Equivalent class axioms are stated in the rest of this Section, class by class.

The following classes are pairwise disjoint:

- AO_Resource
- AO_Object
- AO_Concept
- AO_Agent
- AO_Spatial_Region
- AO_Temporal_Region
- AO_Event

The axiom expressing the above disjointness statement is given by:

DisjointClasses(AO_Resource AO_Object AO_Concept AO_Agent AO_Spatial_Region AO_Temporal_Region AO_Event)

No disjoint union of class expression is part of AO-Cat.

AO_Entity

AO_Entity has as instances all entities in the ARIADNE infrastructure.

Subclass of:	crm:E1_CRM_Entity
Superclass of:	AO_Resource, AO_Object, AO_Concept, AO_Event, AO_Agent, AO_Spatial_Region, AO_Temporal_Region
Domain of:	has_identifier, has_type, has_title, has_description
Range of:	refers_to

AO_Resource

AO_Resource has as instances all archaeological resources that are represented in the AC. These resources are categorised in the following classes (see Figure below):

- AO_Data_Resource**,** representing all types of digital data object described in the Catalogue for discovery, access and integration. Two disjoint classes of data resources are further distinguished:
    - AO_Individual_Data_Resource, including all data resources that are considered as not further decomposed; these include datasets, databases, GIS, and so on.
    - AO_Collection, representing aggregations of data resources considered as wholes, possibly including other collections.
- AO_Service, representing the services owned by the A+ partners and, possibly, made available to other partners within the project on the ARIADNE infrastructure. This class is also suitable for describing new services created within the project.

Subclass of:	AO_Entity
Superclass of:	AO_Data_Resource, AO_Service
Domain of:	was_issued, was_modified, has_publisher, has_contributor, has_creator, has_owner, has_responsible
Range of:	no property

AO_Data_Resource

This class specialises the AO_Resource class and is intended to describe the digital resources of the ARIADNE RI that are data. The type of an instance of this class (e.g., GIS, database, collection and the

like) is specified through the has_type property. This is an abstract class: its instances are all and only the instances of its subclasses.

Subclass of:	AO_Resource, crmpe:PE22_Persistent_Dataset (which is a subclass of crm:E73_Information_Object)
Superclass of:	AO_Collection, AO_Individual_Data_Resource
Domain of:	has_original_id, is_about, has_ARIADNE_subject, has_native_subject, has_derived_subject, has_language, was_created_on, has_landing_page, has_access_policy, has_access_rights, has_extent, has_temporal_coverage, has_spatial_coverage
Range of:	has_part

AO_Individual_Data_Resource

This class is a specialisation of the class AO_Data_Resource, and has as instances data resources that from the ARIADNE RI point of view are atomic, that is no further decomposed. The complementary class is that of AO_Collection (see below), having as instances data resources that are non-atomic, that are composed (of other data resources).

Subclass of:	AO_Data_Resource
Domain of:	no property
Range of:	no property

AO_Document

This class is a specialisation of the class AO_Individual_Data_Resource, and has as instances documents of interest to the ARIADNE RI.

Subclass of:	AO_Individual_Data_Resource
Domain of:	no property
Range of:	no property

AO_Collection

This class is a specialisation of the class AO_Data_Resource, and has as instances collections in the archaeological and archaeological-related domains. An archaeological collection is an aggregation of resources, called the members in the collection, instances of AO_Data_Resource. As a consequence, collections can have other collections as members.

Subclass of:	AO_Data_Resource
Domain of:	has_part
Range of:	no property

AO_Service

This class specialises the AO_Resource class and is intended to describe services in the archaeological domain. A service is "an offer by some actor of their willingness and ability to execute an activity or series of activities upon request" [PEM Specifications 3.1]. Types of services include: stand-alone service, web service, front-office services and back-office service, as specified by the has_type property.

Subclass of:	AO_Resource
Equivalent to:	crmpe:PE8_E-Service, which is a subclass of crm:E7_Activity
Domain of:	is_accessible_at, has_functionality, has_consumed_media, has_produced_media, has_consumed_format, has_produced_format, has_supported_language, has_technical_support
Range of:	no property

AO_Object

This class includes as instances physical objects of interest to the ARIADNE infrastructure.

Subclass of:	AO_Entity, crm:E18_Physical_Thing
Equivalent to:	crmpe:PE8_E-Service (which is a subclass of crm:E7_Activity)
Domain of:	has_time_interval, has_space_region, was_present_at
Range of:	no property

AO_Event

This class includes as instances events of interest to the ARIADNE RI, that is, events that play a role in the ARIADNE AC. An event is a "change of states in cultural, social or physical systems, regardless of scale, brought about by a series or group of coherent physical, cultural, technological or legal phenomena" [CIDOC CRM Specs 6.2].

Subclass of:	AO_Entity, crm:E5_Event
Superclass of:	AO_Activity
Domain of:	occurs_in, happens_at, contains_event
Range of:	contains_event, was_present_at

AO_Activity

This class includes as instances activities of interest to the ARIADNE RI, that is, events that play a role in the ARIADNE AC. Activities are "actions intentionally carried out by actors that result in changes of state in the cultural, social, or physical systems documented" [CIDOC CRM Specs 6.2].

Subclass of:	AO_Event, crm:E7_Activity
Domain of:	no property
Range of:	no property

AO_Agent

"This class comprises people, either individually or in groups, who have the potential to perform intentional actions of kinds for which someone may be held responsible." [CIDOC CRM Specs 6.2].

Subclass of:	AO_Entity, crm:E39_Actor
Superclass of:	AO_Person, AO_Group
Domain of:	has_name, has_agent_identifier, has_email, has_homepage
Range of:	has_publisher, has_contributor, has_owner, has_responsible, has_technical_support

AO_Person

This class comprises individual human beings.

Subclass of:	AO_Agent, crm:E21_Person
Domain of:	has_institution
Range of:	no property

AO_Group

This class comprises “gatherings or organizations that act collectively or in a similar way due to any form of unifying relationship” [CRM Spec 6.2].

Subclass of:	AO_Agent, crm:E74_Group
Domain of:	no property
Range of:	no property

AO_Temporal_Region

This class comprises temporal extents, having a beginning, an end and a non-zero duration. Temporal regions can have one of the following forms:

1. a temporal interval (e.g., from 155 BC to 243 AD). In this case the properties from and until are used to give the boundaries of the temporal interval
2. any period identified by a name (*e.g.*, *“*Neolithic*”*) expressed via the has_period or has_native_period property

- In the case where BC dates have to be supplied, a minus (-) sign could be used as indicated in the expanded Year representation of ISO 8601 (http://www.iso.org/iso/catalogue_detail?csnumber=40874).

- In the case where reduced precision must be applied (e.g. where no day information is available) the respective part could be omitted (according to ISO 8601 reduced precision guidelines).

Subclass of:	AO_Entity, crm:E52_Time-Span
Domain of:	has_period, has_native_period, from, until

Range of:	has_temporal_coverage, happens_at, has_time_interval

AO_Spatial_Region

This class comprises spatial regions having one of the following forms:

1. a point on the surface of the Earth identified by latitude and longitude;
2. a polygon as represented by GIS systems;
3. a rectangular region on the surface of the Earth identified by its four vertices;
4. any region identified by a IRI in a standard gazetteer, such as Geonames for modern places, Pleiades for ancient places.

Each of these forms is modelled by a different sub-class of this class, as detailed in the rest of this Section. In addition, a spatial region may have a name, regardless of any other representation it may have. If the same region is represented in two different ways, then it will be an instance of the two corresponding classes.

Subclass of:	AO_Entity, crm:E53_Place
Superclass of:	AO_Spatial_Region_Point, AO_Spatial_Region_Polygon, AO_Spatial_Region_BBox, AO_Spatial_Region_StdName
Domain of:	has_coordinate_system, has_place_name, has_place_IRI
Range of:	occurs_in, has_spatial_coverage, has_space_region

AO_Spatial_Region_Point

This class comprises spatial regions given as points on the surface of the Earth identified by latitude and longitude.

Subclass of:	AO_Spatial_Region
Domain of:	has_latitude, has_longitude
Range of:	no property

AO_Spatial_Region_Polygon

This class comprises spatial regions given as a polygon as represented by GIS systems.

Subclass of:	AO_Spatial_Region
Domain of:	has_polygonal_representation
Range of:	no property

AO_Spatial_Region_BBox

This class comprises spatial regions given as rectangular regions on the surface of the Earth identified by its four vertices.

Subclass of:	AO_Spatial_Region
Domain of:	has_bounding_box_min_lat, has_bounding_box_min_lon, has_bounding_box_max_lat, has_bounding_box_max_lon
Range of:	no property

AO_Spatial_Region_StdName

This class comprises spatial regions identified by a name expressed via the has_place_name property.

Subclass of:	AO_Spatial_Region
Domain of:	has_place_name
Range of:	no property

AO_Concept

This class comprises terms in thesauri, controlled vocabularies or any other reference resource providing concepts in a domain of interest. Since AO_Concept is equivalent to skos:Concept, all properties defined in SKOS can be used to represent the relevant aspects of a concept. For instance, the properties skos:broader and skos:narrower can be used to model concept taxonomies, while skos:broadMatch, skos:closeMatch, skos:exactMatch, skos:narrowMatch, skos:relatedMatch can be used to specify the kind of mapping leading to the instance of the concept, if any.

Subclass of:	AO_Entity
Equivalent to:	crm:E55 Type, skos:Concept
Domain of:	no property
Range of:	has_type, has_ARIADNE_subject, has_native_subject, has_derived_subject, has_language, has_native_period, has_functionality, has_consumed_media, has_produced_media, has_supported_language

Object property axioms

OWL 2 DL, the most expressive decidable logic for ontology representation in the Semantic Web family, provides the following kinds of axioms on object properties:

- Object Subproperties
- Equivalent Object Properties
- Disjoint Object Properties
- Inverse Object Properties
- Object Property Domain
- Object Property Range
- Functional Object Properties
- Inverse-Functional Object Properties
- Reflexive Object Properties
- Irreflexive Object Properties
- Symmetric Object Properties
- Asymmetric Object Properties
- Transitive Object Properties

All these kinds of axioms will be considered in the rest of this Section. In addition, axioms concerning the cardinality of object properties will be given. These axioms are expressed as subclass or equivalent class axioms on class expressions. For the ease of the reader, cardinality axioms will be spelled in natural language first and then the corresponding axioms will be given.
Notice that the AO-Cat document limits itself to the more common kinds of axioms. In what follows, we will recapitulate these axioms and will provide additional axioms of the above kinds, for each object property of AO-Cat.
In addition, the present document gives axioms capturing mappings from AO-Cat to the CIDOC CRM. The specification of these axioms is derived from the mapping specifications given in the AO-Cat document, such as:

has_publisher → L10i_was_input_of → D7_Digital_Machine_event → P1i_is_domain_of → PC14_carried_out_by → P2_has_range → E39_Actor

where P is an AO-Cat property. This expression is an abbreviation of the following logical sentence:

(forall x,y)	has_publisher(x,y)				
	IMPLIES	(exists wz)	L10i_was_input_of(x,w)	AND	
			D7_Digital_Machine_event(w)	AND	(*)
			P1i_is_domain_of(w,z)	AND	
			PC14_carried_out_by(z)	AND	(*)
			P2_has_range(z,y)	AND	
			E39_Actor(y)		(*)

The assertions in the consequent of the above sentence marked with (*) are redundant because they are implied by domain and range axioms on the involved properties, they are inserted only for convenience.
In general, it is not possible to directly translate the above sentence into an OWL 2 DL axiom. However, it turns out that for any property assertion on property has_publisher:

ObjectPropertyAssertion(has_publisher a b)

it is always possible to find a set of OWL 2 DL assertions that is equivalent to the translation of the above assertion in CRM. This set of assertions is derived by first considering the graph resulting from the translation of the above assertion in RDF, involving almost always unnamed individuals (aka blank nodes), and then transforming the graph into an equivalent set of OWL 2 DL axioms without unnamed individuals, following the method given in [ref OWL 2 DL]. Moreover, these transformations follow a few common patterns, which will be identified in due course and recalled in the sequel.

has_type

This property associates an ARIADNE Entity with a term from a controlled vocabulary.

Every AO_Entity must have a type, it can have any number of types and a type can be associated to any number of AO_Entities. This is captured by making the property mandatory.

Subproperty of:	dct:type
Equivalent to:	crm:P2_has_type
Disjoint from:	all other object properties
Inverse:	is_type_of
Domain:	AO_Entity

Range:	AO_Concept
Mandatory:	yes

has_part

Associates an ARIADNE collection with any AO_Data_Resource resource that the collection contains.

A collection must have a member, can have any number of members, and a data resource can be a member of at most one collection. Moreover, a collection cannot be a member of itself. In other words, every collection is the root of a tree which may be a sub-tree of another collection. This is captured by making the property mandatory, transitive, irreflexive and inverse functional.

Superproperty of:	dct:hasPart, crm:P106_is_composed_of
Disjoint from:	all other object properties
Inverse:	is_part_of
Domain:	AO_Collection
Range:	AO_Data_Resource
Irreflexive:	yes
Transitive:	yes
Mandatory:	yes
Inverse-Functional:	yes

has_publisher

Associates any ARIADNE resource with an agent responsible for making the resource publicly accessible (via download, or API, or other).

An AO_Resource must have a publisher, cannot have more than one, and AO_Agent can be the publisher of any number of resources. This is captured by making the property mandatory and functional.

Subproperty of:	dct:publisher
Disjoint from:	all other object properties
Inverse:	is_publisher_of

Domain:	AO_Resource
Range:	AO_Agent
Mandatory:	yes
Functional:	yes

This property is a shortcut of the fully developed paths:

→L10i_was_input_of→D7_Digital_Machine_event→
P1i_is_domain_of→PC14_carried_out_by→
P2_has_range→E39_Actor

→L10i_was_input_of→D7_Digital_Machine_event→
P1i_is_domain_of→PC14_carried_out_by→
P14.1_in_the_role_of→E55["Publisher"]

→L10i_was_input_of→D7_Digital_Machine_event→
P2_has_type→E55_Type["Ariadne Content Provision"]

which can be abbreviated as:

→ L10i_was_input_of → D7_Digital_Machine_event →
(P2_has_type → E55_Type["Ariadne Content Provision"] |
P1i_is_domain_of → PC14_carried_out_by → (P2_has_range → E39_Actor |
P14.1_in_the_role_of → E55_Type["Publisher"])
)

We have observed that the application of the above path expression to a positive object property assertion on property has_publisher and on the named individuals (i.e., IRIs) a and b:

ObjectPropertyAssertion(has_publisher a b)

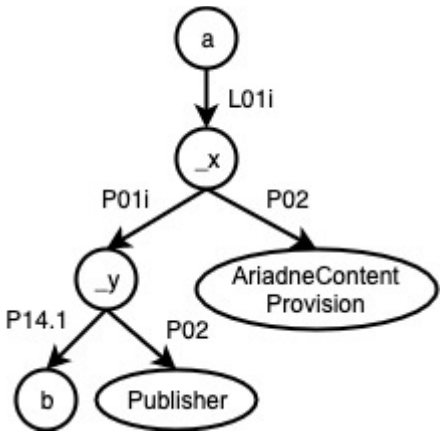
produces the following set S of positive object property assertions (for readability, only the first part of CRM class and property names are given):

S = { ObjectPropertyAssertion(L10i a _x),
ClassAssertion(D7 _x),
ObjectPropertyAssertion(P2 _x ao\type\Ariadne Content Provision),
ObjectPropertyAssertion(P1i _x _y),
ClassAssertion(PC14 _y),
ObjectPropertyAssertion(P14.1 _y ao\type\Publisher),
ObjectPropertyAssertion(P2 _y b) }

where _x and _y are unnamed individuals. Class assertions are redundant and therefore can be in fact omitted from S since they are implied by the axioms on the ranges of properties L10i and P1i. By doing so, the set S' is obtained:

S' = { ObjectPropertyAssertion(L10i a _x),
ObjectPropertyAssertion(P2 _x ao\type\Ariadne Content Provision),
ObjectPropertyAssertion(P1i _x _y),
ObjectPropertyAssertion(P14.1 _y ao\type\Publisher),
ObjectPropertyAssertion(P2 _y b) }

which can be graphically represented as follows:



S' would end up in the A+ Catalog as the CRM translation of the previous object property assertion. The A+ Catalog is technically an OWL 2 DL Ontology and as such it must satisfy, amongst others, the global restriction on unnamed individuals. Happily, S' does satisfy this restriction thanks to its tree structure, therefore the assertions in it can be rephrased as the following equivalent assertion without unnamed individuals:

```
ClassAssertion(
  ObjectSomeValuesFrom( L10i
    ObjectIntersectionOf(
      ObjectHasValue( P2 ao\type\Ariadne Content Provision )
      ObjectSomeValuesFrom( P1i
          ObjectIntersectionOf(
          ObjectHasValue( P14.1 ao\type\Publisher )
          ObjectHasValue( P2 b )
        ))))
  a)
```

This translation into CRM of AO-Cat assertions on property has_publisher follows a pattern that will be called T1.

has_contributor

Associates any ARIADNE resource with an agent primarily responsible for contributing the description of the resource to the ARIADNE Content Cloud.

An ARIADNE Resource must have a contributor, cannot have more than one, and AO_Agent can be the contributor of any number of resources. This is captured by making the property mandatory and functional.

Subproperty of:	dct:contributor
Disjoint from:	all other object properties
Inverse:	is_contributor_of
Domain:	AO_Resource
Range:	AO_Agent
Mandatory:	yes
Functional:	yes

This property is a shortcut of the fully developed path:

→ P94i_was_created_by → E65_Creation →
(P2_has_type → E55_Type["Ariadne Content Creation"] |
P1i_is_domain_of → PC14_carried_out_by → (P2_has_range→E39_Actor |
P14.1_in_the_role_of→E55_Type["Contributor"])
)

which when applied to individuals a and b produces pattern T1, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P94i
    ObjectIntersectionOf(
      ObjectHasValue( P2 ao\type\Ariadne Content Creation )
      ObjectSomeValuesFrom( P1i
          ObjectIntersectionOf(
          ObjectHasValue( P14.1 ao\type\Contributor )
          ObjectHasValue( P2 b )
        ))))
  a)
```

has_creator

Associates any ARIADNE resource with an agent primarily responsible for creating the resource.

An ARIADNE Resource must have a creator, cannot have more than one, and AO_Agent can be the creator of any number of resources. This is captured by making the property mandatory and functional.

Subproperty of:	dct:creator
Disjoint from:	all other object properties
Inverse:	is_creator_of
Domain:	AO_Resource
Range:	AO_Agent
Mandatory:	yes
Functional:	yes

This property is a shortcut of the fully developed path:

→ P94i_was_created_by → E65_Creation →
(P2_has_type → E55_Type["Ariadne Content Creation"] |
P1i_is_domain_of → PC14_carried_out_by → (P2_has_range→E39_Actor |
P14.1_in_the_role_of→E55_Type["Creator"])
)

which when applied to individuals a and b produces pattern T1, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P94i
    ObjectIntersectionOf(
      ObjectHasValue( P2 ao\type\Ariadne Content Creation )
      ObjectSomeValuesFrom( P1i
         ObjectIntersectionOf(
        ObjectHasValue( P14.1 ao\type\Contributor )
        ObjectHasValue( P2 b )
       ))))
  a)
```

has_owner

Associates any ARIADNE resource with the legal owner of the resource, who holds the legal responsibility.

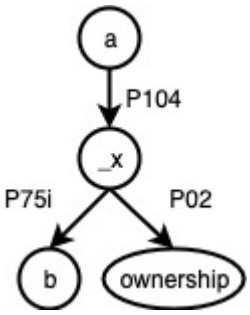
An ARIADNE Resource must have an owner, cannot have more than one, and AO_Agent can be the owner of any number of resources, including zero. This is captured by making the property mandatory and functional.

Disjoint from:	all other object properties
Inverse:	is_owner_of
Domain:	AO_Resource
Range:	AO_Agent
Mandatory:	yes
Functional:	yes

This property is a shortcut of the fully developed path:

→ P104_is_subject_to → E30_Right →
(P2_has_type → E55_Type["ownership"] | P75i_is_possessed_by → E39_Actor))

which when applied to individuals a and b produces a pattern T2 that is a simplified version of T1, graphically given by:



therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P104
    ObjectIntersectionOf(
      ObjectHasValue( P2 ao\type\ownership )
      ObjectHasValue( P75i b )
```

```
    ))
  a)
```

has_responsible

Associates any ARIADNE resource with an agent holding the scientific responsibility for the resource, such as the person or team who conceived the service or gathered the data.

An ARIADNE Resource must have a responsible agent, cannot have more than one, and AO_Agent can be the responsible agent of any number of resources, including zero. This is captured by making the property mandatory and functional.

Disjoint from:	all other object properties
Inverse:	is_responsible_of
Domain:	AO_Resource
Range:	AO_Agent
Mandatory:	yes
Functional:	yes

This property is a shortcut of the fully developed path:

```
→ L10i_was_input_of → D7_Digital_Machine_event →
( P2_has_type → E55_Type["Ariadne Content Provision"] |
        P1i_is_domain_of → PC14_carried_out_by → ( P2_has_range→E39_Actor |
        P14.1_in_the_role_of→E55_Type["Scientific or Technical Responsible"])
)
```

which when applied to individuals a and b produces pattern T1, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( L10i
    ObjectIntersectionOf(
      ObjectHasValue( P2 ao\type\Ariadne Content Provision )
      ObjectSomeValuesFrom( P1i
         ObjectIntersectionOf(
         ObjectHasValue( P14.1 ao\type\Scientific or Technical Responsible )
         ObjectHasValue( P2 b )
        ))))
  a)
```

refers_to

This property is used to associate an AO_Data_Resource to the AO_Entity(ies) to which the resource refers.

A Data Resource can refer to zero, one or more Entities, and an Entity can be referenced by any number of Data Resources.

Subproperty of:	crm:P67_refers_to
Superproperty of:	is_about
Disjoint from:	all other object properties, except its super- or subproperties
Inverse:	is_referenced_by
Domain:	AO_Data_Resource
Range:	AO_Entity
Mandatory:	yes

is_about

This property is a sub-property of refers_to, used to associate an AO_Data_Resource to the primary AO_Concept(s) the resource is about. It is an abstract property that is never directly instantiated, but always through one of its three sub-properties (defined next):

- has_ARIADNE_subject, associating the resource with one of the fundamental archaeological categories defined by ARIADNE
- has_native_subject, associating the resource with an original subject in the providing institution
- has_derived_subject, associating the resource with a subject automatically derived by mapping an original subject to the Getty AAT vocabulary

A Data Resource must be about an AO_Entity, can be about more than one, and AO_Entity can be the subject of any number of Data Resources. This is captured by making the property mandatory.

Subproperty of:	refers_to, crm:P129_is_about, dc:subject
Superproperty of:	has_ARIADNE_subject, has_native_subject, has_derived_subject
Disjoint from:	all other object properties, except its super- or subproperties
Inverse:	is_subject_of

Domain:	AO_Data_Resource
Range:	AO_Entity
Mandatory:	yes

has_ARIADNE_subject

This property associates an AO_Data_Resource with one the fundamental archaeological categories defined by ARIADNE. These are the high level “resource types”, or semantic categories which are used in the ARIADNE portal to filter search results, such as for instance “Site and monuments databases or inventories”, “Event/intervention resources” etc.

A Data Resource must be about an AO_Concept, can be about more than one, and AO_Concept can be the subject of any number of Data Resources. This is captured by making the property mandatory.

Subproperty of:	is_about
Disjoint from:	all other object properties
Inverse:	is_ARIADNE_subject_of
Domain:	AO_Data_Resource
Range:	AO_Concept
Mandatory:	yes

This property is a shortcut of the fully developed path:

→ P129_is_about → E55_Type → P2_has_type → E55_Type["Ariadne Subject"]

We have observed that the application of the above path expression to a positive object property assertion on property has_ARIADNE_subject and on the named individuals (i.e., IRIs) a and b:

ObjectPropertyAssertion(has_ARIADNE_subject a b)

produces the following set S of positive object property assertions (for readability, only the first part of CRM class and property names are given):

S = { ObjectPropertyAssertion(P129 a b),
ObjectPropertyAssertion(P2 b ao\type\Ariadne Subject) }

which we call pattern T3. S does not contain any unnamed individual and therefore when added to an OWL 2 DL ontology does not pose any problem with respect to the global restrictions on the axioms of the ontology.

has_native_subject

This property associates an AO_Data_Resource with an original subject in the providing institution.

A Data Resource must be have a native subject, can have more than one, and a subject can be the subject of any number of Data Resources. This is captured by making the property mandatory.

Subproperty of:	is_about
Disjoint from:	all other object properties
Inverse:	is_native_subject_of
Domain:	AO_Data_Resource
Range:	AO_Concept
Mandatory:	yes

This property has_native_subject is a shortcut of the fully developed path

→ P129_is_about → E55_Type → P2_has_type → E55_Type["Native Subject"]

which when applied to a positive object property assertion on property has_derived_subject and on the named individuals (i.e., IRIs) a and b:

ObjectPropertyAssertion(has_native_subject a b)

produces pattern T3 therefore can be rephrased as the following set S of positive object property assertions (for readability, only the first part of CRM class and property names are given):

S = { ObjectPropertyAssertion(P129 a b),
ObjectPropertyAssertion(P2 b ao\type\Native Subject) }

which does not contain any unnamed individual and therefore when added to an OWL 2 DL ontology does not pose any problem with respect to the global restrictions on the axioms of the ontology.

has_derived_subject

This property associates an AO_Data_Resource with a subject automatically derived by mapping a native subject of the resource to a term in the Getty AAT vocabulary. The native subject may have the SKOS relations (skos:broadMatch, skos:closeMatch,skos:exactMatch,skos:narrowMatch, skos:relatedMatch) to the Getty AAT vocabulary.

A Data Resource must have a derived subject, can have more than one, and a subject can be the derived subject of any number of Data Resources. This is captured by making the property mandatory.

Subproperty of:	is_about
Disjoint from:	all other object properties
Inverse:	is_derived_subject_of
Domain:	AO_Data_Resource
Range:	AO_Concept (Getty AAT Term)
Mandatory:	yes

This property is a shortcut of the fully developed path

→ P129_is_about → E55_Type → P2_has_type → E55_Type["Derived Subject"]

which when applied to a positive object property assertion on property has_derived_subject and on the named individuals (i.e., IRIs) a and b:

ObjectPropertyAssertion(has_derived_subject a b)

produces pattern T3 therefore can be rephrased as the following set S of positive object property assertions (for readability, only the first part of CRM class and property names are given):

S = { ObjectPropertyAssertion(P129 a b),
ObjectPropertyAssertion(P2 b ao\type\Derived Subject) }

which does not contain any unnamed individual and therefore when added to an OWL 2 DL ontology does not pose any problem with respect to the global restrictions on the axioms of the ontology.

has_language

This property is used to associate a data resource with the language(s) of the resource, specified according to the vocabulary lexvo.org.

A Data Resource must have a language, can have more than one, and a language can be the language of any number of Data Resources. This is captured by making the property mandatory.

Subproperty of:	dc:language
Equivalent to:	crm:P72_has_language

Disjoint from:	all other object properties
Inverse:	is_language_of
Domain:	AO_Data_Resource
Range:	AO_Concept (linked to lexvo.org)
Mandatory:	yes

has_landing_page

This property is used to specify the original landing page of an AO_Data_Resource, if any.

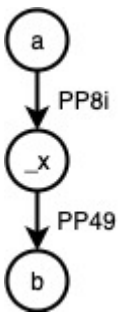
A data resource may have at most one landing page and a web page can be the landing page of any number of data resources, including zero. This is captured by making this property functional.

Disjoint from:	all other object properties
Inverse:	is_landing_page_of
Domain:	AO_Data_Resource
Range:	rdfs:Resource
Functional:	yes

This property is a shortcut of the fully developed path:

→ PP8i_is_dataset_hosted_by → PE15_Data_E-Service → PP49_provides_access_point → PE29_Access_Point

which when applied to individuals a and b produces a pattern T4 that is another simplified version of T1, graphically given by:



therefore can be rephrased as the following equivalent assertion:

ClassAssertion(
  ObjectSomeValuesFrom(PP8i
    ObjectHasValue(PP49 b))
  a)

has_access_policy

This property is used to specify the URI to the statement of policy (typically, on an organisation's website) for the data resource.

A data resource may have at most one access policy and a web page can give the access policy of any number of data resources. This is captured by making this property functional.

Disjoint from:	all other object properties
Inverse:	is_access_policy_of
Domain:	AO_Data_Resource
Range:	rdfs:Resource
Functional:	yes

This property is a shortcut of the fully developed path

→ P104_is_subject_to → E30_Right → P2_has_type → E55_Type[“Access Policy”]

which when applied to individuals a and b produces pattern T3, therefore can be rephrased as the following equivalent set of assertions:

S = { ObjectPropertyAssertion(P104 a b),
  ObjectPropertyAssertion(P2 b ao\type\Access Policy) }

which does not contain any unnamed individual and therefore when added to an OWL 2 DL ontology does not pose any problem with respect to the global restrictions on the axioms of the ontology.

has_temporal_coverage

This property associates a data resource with the temporal region covered by the content of the resource.

A data resource may have at most one temporal coverage and a temporal region can give the coverage of any number of data resources. This is captured by making this property functional.

Subproperty of:	dct:temporal

Equivalent to:	crm:P129_is_about with a restricted range E4_Period
Disjoint from:	all other object properties
Inverse:	is_temporal_coverage_of
Domain:	AO_Data_Resource
Range:	AO_Temporal_Region
Functional:	yes

occurs_in

This property associates an ARIADNE Event with the spatial region in which the event occurred.

An event occurs in exactly one spatial region and a spatial region may be the place of occurrence of any number of events. This is captured by making the property mandatory and functional.

Equivalent to:	crm:P7_took_place_at
Disjoint from:	all other object properties
Inverse:	is_spatial_region_of
Domain:	AO_Event
Range:	AO_Spatial_Region
Mandatory:	yes
Functional:	yes

happens_at

This property associates an ARIADNE Event with the temporal region at which the event occurred.

An event may have at most one temporal coverage and a temporal region can be the occurrence time of any number of events. This is captured by making this property functional.

Equivalent to:	crm:P4_has_time-span
Disjoint from:	all other object properties
Inverse:	is_temporal_region_of

Domain:	AO_Event
Range:	AO_Temporal_Region
Functional:	yes

contains_event

This property associates an ARIADNE Event with another ARIADNE Event that is part of it.

An event may contain zero, one or more events, and may be contained in any number of events, including zero. An event cannot contain itself, either directly or through some other events. In other words, every event is the root of a tree which may be a sub-tree of another event. This is captured by making the property transitive, irreflexive and inverse functional.

Equivalent to:	crm:P9_consists_of
Disjoint from:	all other object properties
Inverse:	is_contained_in_event
Domain:	AO_Event
Range:	AO_Event
Irreflexive:	yes
Transitive:	yes
Inverse-Functional:	yes

has_period

This property associates an AO_Temporal_Region with a temporal period defined in periodO and identified by an IRI, so it ranges over web resources.

A temporal region can have at most a corresponding periodO period and a periodO period may be associated to any number of temporal regions. This is captured by making this property functional.

Disjoint from:	all other object properties
Inverse:	is_period_of
Domain:	AO_Temporal_Region

Range:	rdfs:Resource
Functional:	yes

This property is a shortcut of the fully developed path:

→ P2_has_type → E55_Type → P2_has_type → E55_Type[“PeriodO”]

which when applied to individuals a and b produces pattern T3, therefore can be rephrased as the following equivalent set of assertions:

S = { ObjectPropertyAssertion(P2 a b),
ObjectPropertyAssertion(P2 b ao\type\PeriodO) }

which does not contain any unnamed individual and therefore when added to an OWL 2 DL ontology does not pose a problem with respect to the global restrictions on the axioms of the ontology.

has_native_period

This property associates an AO_Temporal_Region with a concept representing a period in some local vocabulary of the provider.

A temporal region can have any number of corresponding native periods, including zero, and a native period can be associated to any number of temporal regions.

Disjoint from:	all other object properties
Inverse:	is_native_period_of
Domain:	AO_Temporal_Region
Range:	AO_Concept

This property is a shortcut of the fully developed path:

→ P2_has_type → E55_Type → P2_has_type → E55_Type[“Native Period”]

which when applied to individuals a and b produces pattern T3, therefore can be rephrased as the following equivalent set of assertions:

S = { ObjectPropertyAssertion(P2 a b),
ObjectPropertyAssertion(P2 b ao\type\Native Period) }

which does not contain any unnamed individual and therefore when added to an OWL 2 DL ontology does not pose a problem with respect to the global restrictions on the axioms of the ontology.

has_spatial_coverage

This property associates a data resource with the spatial region covered by the content of the resource.

A data resource must have exactly one spatial coverage and a spatial region may be the coverage of any number of data resources, including zero. This is captured by making this property mandatory and functional.

Subproperty of:	dc:spatial
Equivalent to:	crm:P129_is_about
Disjoint from:	all other object properties
Inverse:	is_spatial_coverage_of
Domain:	AO_Data_Resource
Range:	AO_Spatial_Region
Mandatory:	yes
Functional:	yes

is_accessible_at

This property associates a service with an IRI where the service is accessible. If the service is a web service, this IRI is the actual access point of the service. Otherwise, it is the IRI of a resource describing how the service can be accessed.

A service must have exactly one access point. A web resource may be the access point of any number of services, including zero. This is captured by making this property mandatory and functional.

Equivalent to:	crmpe:PP28_has_designated_access_point
Disjoint from:	all other object properties
Inverse:	is_access_point_of
Domain:	AO_Service
Range:	rdfs:Resource
Mandatory:	yes
Functional:	yes

has_functionality

This property associates a service with its functionality, expressed using a controlled vocabulary.

A service has at least one and possibly many functionalities, and a functionality is associated with any number of services. This is captured by making this property mandatory.

Equivalent to:	crmpe:PP45_has_competency
Disjoint from:	all other object properties
Inverse:	is_functionality_of
Domain:	AO_Service
Range:	AO_Concept
Mandatory:	yes

has_consumed_media

This property associates a service with the media type(s) handled by the service. The list of possible mediaType is open.

A service is associated with at least one, possibly many, consumed media types and a media type can be associated with any number of services. This is captured by making this property mandatory.

Equivalent to:	crm:P125_used_object_of_type
Disjoint from:	all other object properties
Inverse:	is_consumed_media_of
Domain:	AO_Service
Range:	AO_Concept
Mandatory:	yes

has_produced_media

This property specifies the media types of the objects created by the service. The list of possible mediaType is open.

A service is associated with at least one, possibly many, produced media types and a media type can be associated with any number of services. This is captured by making this property mandatory.

Disjoint from:	all other object properties
Inverse:	is_produced_media_of
Domain:	AO_Service
Range:	AO_Concept
Mandatory:	yes

This property is a shortcut of the fully developed path:

→ L11_had_output- → D1_Digital_Object- → P2_has_type → E55_Type → P2_has_type → E55_Type[“media”]

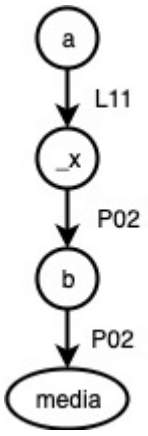
We have observed that the application of the above path expression to a positive object property assertion on property has_produced_media and on the named individuals (i.e., IRIs) a and b:

ObjectPropertyAssertion(has_produced_media a b)

produces the following set S of positive object property assertions (for readability, only the first part of CRM class and property names are given):

S = { ObjectPropertyAssertion(L11 a _x),
ObjectPropertyAssertion(P2 _x b),
ObjectPropertyAssertion(P2 b ao\type\media) }

which has a pattern T5 that is a combination of pattern T3 with a simplified version of T1, graphically given by:



S satisfies the global restriction on unnamed individuals and its first two assertions can be replaced by the equivalent assertion:

ClassAssertion(ObjectSomeValuesFrom(L11 ObjectSomeValuesFrom(P2 b)) a)

has_consumed_format

This property specifies the MIME type of the objects handled by the service.

An object has exactly one MIME type, while a MIME type can be the type of any number of objects. This is captured by making this property mandatory and functional.

Disjoint from:	all other object properties
Inverse:	is_consumed_format_of
Domain:	AO_Service
Range:	rdfs:Resource (a IRI identifying a MIME type)
Mandatory:	yes
Functional:	yes

This property is a shortcut of the fully developed path:

→ crmdig:L10_had_input → crmdig:D1_Digital_Object- → crm:P2_has_type → crm:E55_Type → crm:P2_has_type → crm:E55_Type[“format”]

which when applied to individuals a and b produces pattern T5, therefore can be rephrased as the following equivalent set of assertions:

S = { ClassAssertion(ObjectSomeValuesFrom(L10 ObjectSomeValuesFrom(P2 b)) a),
ObjectPropertyAssertion(P2 b ao\type\format) }

has_produced_format

This property specifies the MIME type of the objects created by the service.

An object has exactly one MIME type, while a MIME type can be the type of any number of objects. This is captured by making this property mandatory and functional.

Disjoint from:	all other object properties
Inverse:	is_produced_format_of

Domain:	AO_Service
Range:	rdfs:Resource (a IRI identifying a MIME type)
Mandatory:	yes
Functional:	yes

This property is a shortcut of the fully developed path:

→ crmdig:L11_had_output → crmdig:D1_Digital_Object → crm:P2_has_type → crm:E55_Type → cm:P2_has_type → crm:E55_Type[“format”]

which when applied to individuals a and b produces pattern T5, therefore can be rephrased as the following equivalent set of assertions:

S = { ClassAssertion(ObjectSomeValuesFrom(L11 ObjectSomeValuesFrom(P2 b)) a),
ObjectPropertyAssertion(P2 b ao\type\format) }

has_supported_language

This property specifies the languages supported by the service, encoded according with ISO 639 standard (ISO 639-1:2002).

A service may be supported by any number of languages including zero and vice-versa.

Subproperty of:	dc:languageFunctional
Equivalent to:	crm:P72_has_language
Disjoint from:	all other object properties
Inverse:	is_supported_language_of
Domain:	AO_Service
Range:	AO_Concept (linked to lexvo.org)

has_technical_support

This property specifies the agent offering technical support on the service.

A service has exactly one technical supporting agent, while an agent can give support to any number of services, including zero. This is captured by making this property mandatory and functional.

Subproperty of:	crmpe:PP2_providedoffered_by
Disjoint from:	all other object properties
Inverse:	is_technical_support_of
Domain:	AO_Service
Range:	AO_Agent
Mandatory:	yes
Functional:	yes

has_time_interval

This property connects an object to a temporal region relevant for that object.

An object is connected to at most one temporal region, while a temporal region may have any number of connecting objects, including zero. This is captured by making this property functional.

Disjoint from:	all other object properties
Inverse:	is_time_interval_of
Domain:	AO_Object
Range:	AO_Temporal_Region
Functional:	yes

This property is a shortcut of the fully developed path:

→ crm:P140i_was_attributed_by → crm:E13_Attribute_Assignment → crm:P141_assigned → crm:E52_Time-Span

which when applied to individuals a and b has pattern T4 therefore can be rephrased as the following equivalent assertion:

ClassAssertion(ObjectSomeValuesFrom(P140i ObjectHasValue(P141 b)) a)

has_space_region

This property connects an object to a spatial region relevant for that object.

An object is connected to at most one spatial region, while a spatial region may have any number of connecting objects, including zero. This is captured by making this property functional.

Subproperty of:	crm:P53_has_former_or_current_location
Disjoint from:	all other object properties
Inverse:	is_space_region_of
Domain:	AO_Object
Range:	AO_Spatial_Region
Functional:	yes

was_present_at

This property connects an object to an event in which the object plays a relevant role.

An object may be connected to any number of events, including zero. An event may have any number of connecting objects, including zero.

Subproperty of:	crm:P12i_was_present_at
Disjoint from:	all other object properties
Inverse:	has_present_object
Domain:	AO_Object
Range:	AO_Event

has_homepage

This property is used to specify the website homepage of the person, institution or organisation that is part of the ARIADNE Catalogue.

A person has at most one homepage while a web page can be the homepage of at most one person. This is captured by making the property functional and inverse functional.

Subproperty of:	foaf:homepage
Disjoint from:	all other object properties

Inverse:	is_homepage_of
Domain:	Agent
Range:	rdfs:Resource
Functional:	yes
Inverse-Functional:	yes

This property is a shortcut of the fully developed path:

→ crm:P76_has_contact_point → E51_Contact_Point → P2_has_type → E55_Type[“homepage”]

which when applied to individuals a and b produces pattern T3, therefore can be rephrased as the following equivalent set of assertions:

S = { ObjectPropertyAssertion(P76 a b),
ObjectPropertyAssertion(P2 b ao\type\homepage) }

has_place_IRI

This property associates a spatial region with an IRI or other unique identifier from a standard gazetteer (e.g. Geonames for modern places, Pleiades for ancient places) used to refer to that spatial region.

A spatial region may have at most one IRI from a standard gazetteer, and an IRI from a standard gazetteer may be associated with at most one spatial region. This is captured by making this property functional and inverse functional.

Disjoint from:	all other object properties
Inverse:	is_IRI_of_place
Domain:	AO_Spatial_Region
Range:	rdfs:Resource
Functional:	yes
Inverse-Functional:	yes

This property is a shortcut of the fully developed path:

→ P87_is_identified_by → E44_Place_Appellation → P2_has_type → E55_Type[“gazetteer name”]

which when applied to individuals a and b produces pattern T3, therefore can be rephrased as the following equivalent set of assertions:

S = { ObjectPropertyAssertion(P87 a b),
ObjectPropertyAssertion(P2 b ao\type\gazetteer name) }

Data property axioms

OWL 2 DL provides the following kinds of axioms on data properties:

- Data Subproperties
- Equivalent Data Properties
- Disjoint Data Properties
- Data Property Domain
- Data Property Range
- Functional Data Properties

Similarly to object properties, all data property axioms of the above kinds are given in the rest of this section for every data property of AO-Cat. Mappings to CRM will be treated in the same way they are treated for object properties.

has_identifier

This property associates an ARIADNE Entity with an identifier for that resource in some namespace other than the ARIADNE namespace (where the primary identifier of the resource belongs) and other than the provider's namespace (which is associated to data resources by property has_original_id). The range of this property is xsd:string for generality.

An ARIADNE Entity may have at most one of these identifiers, thus the property is functional.

Superproperty of:	dct:identifier
Disjoint from:	all other object properties
Domain:	AO_Entity
Range:	xsd:string
Functional:	yes

This property is a shortcut of the fully developed path:

crm:P1_is_identified_by → crm:E42_Identifier → rdfs:label → rdfs:Literal

which when applied to individual a and literal i in

ObjectHasValue(has_identifier a i)

produces pattern T4, therefore can be rephrased as the following equivalent assertion:

ClassAssertion(
  ObjectSomeValuesFrom(P1
    DataHasValue(rdfs:label i))
  a)

has_title

This property associates an ARIADNE Entity with one title identifying the resource in a specific language.

An ARIADNE Entity must have at least one title thus the property is mandatory.

An ARIADNE Entity can have more than one title, but in that case the different titles must be in different languages. The uniqueness of the title for a language cannot be captured because the language of a title is not explicitly stated via an AO-Cat property.

Subproperty of:	dct:title
Disjoint from:	all other object properties
Domain:	AO_Entity
Range:	xsd:string
Mandatory:	yes

This property is a shortcut of the fully developed path:

crm:P102_has_title → crm:E35_Title → rdfs:label → rdfs:Literal

which when applied to individual a and literal i in

DataHasValue(has_title a i)

produces pattern T4, therefore can be rephrased as the following equivalent assertion:

ClassAssertion(
  ObjectSomeValuesFrom(P102

DataHasValue(rdfs:label i))
a)

has_description

This property associates an ARIADNE Entity with a free text description of the resource.

An ARIADNE Entity may have at most one description, thus the property is functional.

Subproperty of:	dct:description, rdfs:comment, crm:P3_has_note
Disjoint from:	all other object properties
Domain:	AO_Entity
Range:	xsd:string
Functional:	yes

was_issued

This property associates an ARIADNE resource with the date of formal issuance (e.g., publication) of the resource by the publisher.

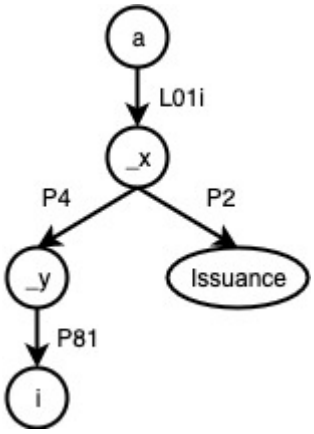
An ARIADNE resource has exactly one date of formal issuance, thus the property is mandatory and functional.

Disjoint from:	all other object properties
Domain:	AO_Resource
Range:	xsd:dateTime
Mandatory:	yes
Functional:	yes

This property is a shortcut of the fully developed path:

→ L10i_was_input_of → D7_Digital_Machine_event →
(P4_has_time-span → E52_Time-Span → P81_ongoing_throughout → xsd:dateTime
| P2_has_type → E55 ["Issuance"])

which when applied to individual a and literal i produces a pattern T6 that is a simplified version of T1, graphically given by:



therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( L10i
    ObjectIntersectionOf(
      ObjectHasValue( P2 ao\type\Issuance )
      ObjectSomeValuesFrom( P4
          DataHasValue( P81 i )
        )))
  a)
```

was_modified

This property associates an ARIADNE resource with the most recent date on which the resource was modified by the publisher.

An ARIADNE resource must be associated with exactly one modification date, thus the property is mandatory and functional.

Disjoint from:	all other object properties
Domain:	AO_Resource
Range:	xsd:dateTime
Mandatory:	yes
Functional:	yes

It is assumed that every modification is a new creation that is a Digital Machine Event, and that the resource is the resulting output of this Digital Machine Event. Therefore, this property is interpreted as a shortcut of the fully developed path:

→ L11i_was_output_of → D7_Digital_Machine_Event →
(P4_has_time-span → E52_Time-Span → P82_at_some_time_within → xsd:date | P2_has_type → E55_Type["Modification"])

which when applied to individual a and literal i produces pattern T6, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( L11i
    ObjectIntersectionOf(
      ObjectHasValue( P2 ao\type\Modification )
      ObjectSomeValuesFrom( P4
          DataHasValue( P82 i )
        )))
  a)
```

has_original_id

This property associates a data resource with the local identifier of the resource supplied by the content provider.

A data resource must have exactly one local identifier, thus the property is mandatory and functional.

Disjoint from:	all other object properties
Domain:	AO_Data_Resource
Range:	xsd:string
Mandatory:	yes
Functional:	yes

This property is a shortcut of the fully developed path:

→ crm:P1_is_identified_by → crm:E42_Identifier →
(rdfs:label → rdfs:Literal | crm:P2_has_type → crm:E55_Type["originalID"])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P1
    DataIntersectionOf(
      ObjectHasValue( P2 ao\type\originalID )
      DataHasValue( rdfs:label i )
        ))
  a)
```

was_created_on

This property is used to specify the creation date of the AO_Data_Resource. This is the date when the AO_Data_Resource was first made available online by the Publisher, (not the date of the fieldwork or laboratory analysis for instance). In some cases (e.g. where data or metadata did not already exist online) it will be the same date as the creation of the metadata for ARIADNE.

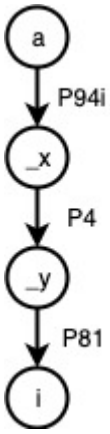
A data resource must have exactly one creation date, thus the property is mandatory and functional.

Subproperty of:	dct:created
Disjoint from:	all other object properties
Domain:	AO_Data_Resource
Range:	xsd:dateTime
Mandatory:	yes
Functional:	yes

This property is a shortcut of the fully developed path:

→ P94i_was_created_by → E65_Creation → P4_has_time-span → E52_Time-Span → P81_ongoing_throughout → xsd:date

which when applied to individual a and literal i produces a pattern (T7) that is a simplification of T1 graphically given by:



therefore can be rephrased as the equivalent assertion:

ClassAssertion(
  ObjectSomeValuesFrom(P94i ObjectSomeValuesFrom(P4 DataHasValue(P81 i))) a)

has_access_rights

This property gives information about who can access a data resource or indicates the data resource's security status.

It is mandatory but not necessarily functional, as a resource may have several pieces of access right information with which it is associated.

Subproperty of:	dct:accessRights
Disjoint from:	all other object properties
Domain:	AO_Data_Resource
Range:	xsd:string
Mandatory:	yes

This property is a shortcut of the fully developed path;

→ P104_is_subject_to → E30_Right →
(rdfs:label → rdfs:Literal | P2_has_type → E55_Type[“Access Rights”])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

ClassAssertion(
  ObjectSomeValuesFrom(P104
    ObjectIntersectionOf(

ObjectHasValue(P2 ao\type\Access Rights)
DataHasValue(rdfs:label i)
))
a)

has_extent

This property is used to specify the size of the AO_Data_Resource (i.e., number of members in a collection, number of records in a dataset, etc.)

A data resource has at most one extent, thus the property is functional.

Superproperty of:	dct:extent
Equivalent to:	crm:P43_has_dimension
Disjoint from:	all other object properties
Domain:	AO_Data_Resource
Range:	xsd:string
Functional:	yes

has_name

This property is used to specify the name of an agent, which may be a person, institution or organisation that is part of the ARIADNE Catalogue.

An agent has exactly one name, thus the property is mandatory and functional.

Subproperty of:	foaf:name
Disjoint from:	all other object properties
Domain:	AO_Agent
Range:	xsd:string
Mandatory:	yes
Functional:	yes

This property is a shortcut of the fully developed path:

→ P1_is_identified_by → E41_Appellation →

(rdfs:label → rdfs:Literal | P2_has_type → E55_Type['Name'])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P1
    ObjectIntersectionOf(
      ObjectHasValue( P2 ao\type\Name )
      DataHasValue( rdfs:label i )
        ))
  a)
```

has_agent_identifier

This property is used to specify the identifier of an agent.

In order to specify the authority that provides the identifier, it is recommended to create a sub-property relative to the authority.

An agent may have any number of identifiers, including zero.

Subproperty of:	foaf:name
Disjoint from:	all other object properties
Domain:	Agent
Range:	xsd:string

This property is a shortcut of the fully developed path

→ P1_is_identified_by → E41_Appellation → rdfs:label → rdfs:Literal

which when applied to individual a and literal i produces pattern T4, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P1
    DataHasValue( rdfs:label i ))
  a)
```

Any sub-property of this property defined for the purpose of specifying the authority providing the identifier would be a shortcut of the fully developed path

→ P1_is_identified_by → E41_Appellation →
(rdfs:label → rdfs:Literal | P2_has_type → E55_Type[authority-iri])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P1
    ObjectIntersectionOf(
      ObjectHasValue( P2 authority-iri )
      DataHasValue( rdfs:label i )
        ))
  a)
```

has_email

This property is used to specify the email address of an agent.

An agent may have zero, one or more than one email addresses.

Subproperty of:	foaf:mbox
Disjoint from:	all other object properties
Domain:	Agent
Range:	xsd:string

This property is a shortcut of the fully developed path:

→ crm:P76_has_contact_point → E41_Appellation →
(rdfs:label → rdfs:Literal | P2_has_type → E55_Type[“Email”])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P76
    ObjectIntersectionOf(
      ObjectHasValue( P2 ao\type\Email )
      DataHasValue( rdfs:label i )
        ))
  a)
```

from

This property is used to indicate the starting date of an AO_Temporal_Region in ISO 8601 format.
A temporal region has at most a starting point, thus the property is functional.

Equivalent to:	crm:P79_beginning_is_qualified_by
Disjoint from:	all other object properties
Domain:	AO_Temporal_Region
Range:	xsd:dateTime
Functional:	yes

until

This property is used to indicate the ending date of an AO_Temporal_Region in ISO 8601 format.

A temporal region has at most an ending point, thus the property is functional.

Equivalent to:	crm:P80_end_is_qualified_by
Disjoint from:	all other object properties
Domain:	AO_Temporal_Region
Range:	xsd:dateTime
Functional:	yes

has_place_name

This property is used to associate a name to a SpatialRegion.

A region must have a place name and may have more than one. Thus the property is mandatory.

Subproperty of:	P87_is_identified_by
Disjoint from:	all other object properties
Domain:	AO_Spatial_Region
Range:	xsd:string

Mandatory:	yes

has_coordinate_system

This property indicates the system used to encode the coordinates of a Spatial Region.

A spatial region has at most one coordinate system, thus the property is functional.

Disjoint from:	all other object properties
Domain:	AO_Spatial_Region
Range:	xsd:string
Functional:	yes

This property is a shortcut of the fully developed path:

→ P87_is_identified_by → E41_Appellation →
(rdfs:label → rdfs:Literal | P2_has_type → E55_Type[*coordinate-system-iri*])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

ClassAssertion(
  ObjectSomeValuesFrom(P87
    ObjectIntersectionOf(
      ObjectHasValue(P2 *coordinate-system-iri*)
      DataHasValue(rdfs:label i)
        ))
  a)

has_latitude

This property indicates the latitude value of the coordinates of an AO_Spatial_Region_Point.

A spatial point has at most one latitude value, thus the property is functional.

Subproperty of:	geo:lat
Disjoint from:	all other object properties
Domain:	AO_Spatial_Region_Point

Range:	xsd:decimal
Functional:	yes

This property is a shortcut of the fully developed path

→ P87_is_identified_by → E41_Appellation →
(rdfs:label → rdfs:Literal | P2_has_type → E55_Type[aocat\type\latitude])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P87
    ObjectIntersectionOf(
      ObjectHasValue( P2 aocat\type\latitude )
      DataHasValue( rdfs:label i )
        ))
  a)
```

has_longitude

This property indicates the latitude value of the coordinates of an AO_Spatial_Region_Point.

A spatial point has at most one latitude value, thus the property is functional.

Subproperty of:	geo:lon
Disjoint from:	all other object properties
Domain:	AO_Spatial_Region_Point
Range:	xsd:decimal
Functional:	yes

This property is a shortcut of the fully developed path:

→ P87_is_identified_by → E41_Appellation →
(rdfs:value → rdfs:Literal | P2_has_type → E55_Type[aocat\type\longitude])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
```

```
  ObjectSomeValuesFrom( P87
    ObjectIntersectionOf(
      ObjectHasValue( P2 aocat\type\longitude )
      DataHasValue( rdfs:value i )
        ))
  a)
```

has_bounding_box_min_lat

This property indicates the minimum latitude value of the bounding box area defining an AO_Spatial_Region_BBox.

A bounding box has at most one minimum latitude value, thus the property is functional.

Disjoint from:	all other object properties
Domain:	AO_Spatial_Region_BBox
Range:	xsd:decimal
Functional:	yes

This property is a shortcut of the fully developed path:

→ P87_is_identified_by → E41_Appellation →
(rdfs:value → rdfs:Literal | P2_has_type → E55_Type[aocat\type\boundingBoxMinLat])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P87
    ObjectIntersectionOf(
      ObjectHasValue( P2 aocat\type\boundingBoxMinLat )
      DataHasValue( rdfs:value i )
        ))
  a)
```

has_bounding_box_min_lon

This property indicates the minimum longitude value of the bounding box area defining an AO_Spatial_Region_BBox.

A bounding box has at most one minimum longitude value, thus the property is functional.

Disjoint from:	all other object properties
Domain:	AO_Spatial_Region_BBox
Range:	xsd:decimal
Functional:	yes

This property is a shortcut of the fully developed path:

→ P87_is_identified_by → E41_Appellation →
(rdfs:value → rdfs:Literal | P2_has_type → E55_Type[aocat\type\boundingBoxMinLon])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P87
    ObjectIntersectionOf(
      ObjectHasValue( P2 aocat\type\boundingBoxMinLon )
      DataHasValue( rdfs:value i )
        ))
  a)
```

has_bounding_box_max_lat

This property indicates the maximum latitude value of the bounding box area defining an AO_Spatial_Region_BBox.

A bounding box has at most one maximum latitude value, thus the property is functional.

Disjoint from:	all other object properties
Domain:	AO_Spatial_Region_BBox
Range:	xsd:decimal
Functional:	yes

This property is a shortcut of the fully developed path:

→ P87_is_identified_by → E41_Appellation →
(rdfs:value → rdfs:Literal | P2_has_type → E55_Type[“boundingBoxMaxLat”])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P87
    ObjectIntersectionOf(
      ObjectHasValue( P2 aocat\type\boundingBoxMaxLat )
      DataHasValue( rdfs:value i )
        ))
  a)
```

has_bounding_box_max_lon

This property indicates the maximum longitude value of the bounding box area defining an AO_Spatial_Region_BBox.

A bounding box has at most one maximum longitude value, thus the property is functional.

Disjoint from:	all other object properties
Domain:	AO_Spatial_Region_BBox
Range:	xsd:decimal
Functional:	yes

This property is a shortcut of the fully developed path:

→ P87_is_identified_by → E41_Appellation →
(rdfs:value → rdfs:Literal | P2_has_type → E55_Type[aocat\type\boundingBoxMaxLon])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P87
    ObjectIntersectionOf(
      ObjectHasValue( P2 aocat\type\boundingBoxMaxLon )
      DataHasValue( rdfs:value i )
        ))
  a)
```

has_polygonal_representation

This property associates a polygonal spatial region with an XML representation of the polygon, typically handled by a GIS.

A polygonal spatial region has at most one representation, thus the property is functional.

Disjoint from:	all other object properties
Domain:	AO_Spatial_Region_Polygon
Range:	rdfs:XMLLiteral
Functional:	yes

This property is a shortcut of the fully developed path:

→ P87_is_identified_by → E41_Appellation →
(rdfs:value → rdfs:Literal | P2_has_type → E55_Type[aocat\type\polygon])

which when applied to individual a and literal i produces pattern T2, therefore can be rephrased as the following equivalent assertion:

```
ClassAssertion(
  ObjectSomeValuesFrom( P87
    ObjectIntersectionOf(
      ObjectHasValue( P2 aocat\type\polygon )
      DataHasValue( rdfs:value i )
        ))
  a)
```

has_institution

This property specifies the institution where a person agent belongs.

A person agent belongs to at most one institution, thus the property is functional.

Subproperty of:	foaf:member
Disjoint from:	all other object properties
Domain:	AO_Person
Range:	xsd:string

Functional:	yes