



Data Aggregation Pipeline: User Guide

Alessi Bardi, Ceri Binding, Achille Felicetti, Carlo Meghini, Julian Richards, Maria Theodoridou, Vangelis Kritsotakis

Version: 2.4

Revision history

Date	Version	Author(s)	Description
25/10/2019	1.1	J. Richards, C. Meghini, A. Bardi, A. Felicetti, M. Theodoridou, C. Binding	First draft
20/12/2019	1.2	J. Richards, C. Meghini, A. Bardi, A. Felicetti, M. Theodoridou, C. Binding	Pre-publication revisions
7/1/2020	1.3	J. Richards	Final copy edit
17/1/2020	1.4	J. Richards & C. Meghini	Addition of Appendix 1
2/2/2020	1.5	C. Meghini	Aligned with modified AO-Cat
9/3/2020	1.6	A.Bardi	Added description of the flow to include date normalization and Timespans
27/04/2020	1.7	J.Richards	Aligned names for ARIADNE_subjects with the agreed terms
18/05/2020	1.8	A.Bardi	Section 4 updated: mapping to PeriodO is not required. Removed option to map normalized temporal coverages separately: TimeSpans will produce the modified XML records as output.
26/05/2020	1.9	J.Richards	Section 4.1 updated with link to CNR partner data collection questionnaire
21/11/2020	2.0	J.Richards	Updated to align with AO-Cat v.1.1; ARIADNE_subjects updated; addition of section 3.2.1 on spatial precision. Updated Section 6 with note about years CE.
20/1/2021	2.1	J.Richards	Added Activity Dash Workflow, and added advice re: data provider names
2/3/2021	2.2	V.Kritsotakis	Added User Guidelines for Activity Dash

3/3/2022	2.3	J.Richards & C. Meghini	Changed the range of from and until properties to xsd:gYear to make them compliant with what we get from PeriodO.
30/5/2022	2.4	J.Richards & C. Meghini	Added 'has_country_code' as a mandatory property for spatial regions

Table of Contents

1. Introduction	6
1.1. Related documents	6
1.2. Structure of the document	6
2. Background: Why are we doing this?	7
3. Matching the requirements: introduction to the AO-Cat	10
3.1. Objects	14
3.2. Defining Space	14
3.2.1. Defining spatial precision	15
3.3. Defining time	16
3.4. The ARIADNE fundamental categories - archaeological data types	17
3.5. Events and activities	18
4. Using the Activity Dash tool	19
4.1. Getting Started	19
4.2. Accessing Project	20
4.3. Accessing Workflow	20
4.4. Explaining the Workflow View	21
4.5. Applying Changes on the Workflow	23
4.6. Providing Feedback	25
4.7. View and respond to messages or conversations	28
4.8. More Detailed Instructions	31
5. Supplying Data to CNR	32
5.1. In a nutshell	32
5.2. Metadata ingestion	33
5.3. Metadata Transformation	34
5.4. Metadata Enrichment	34
5.5. Check your content!	35
5.6. Frequency of updates	35
6. Using the 3M Mapping tool	36
6.1. Quick Start Guide	37
6.2. Basic Interface Operation	37
6.2.1. Login	37
6.2.2. Control toolbar and Mapping list	39
6.2.3. Control toolbar – More/Copy XML/Rights	39
6.2.4. Search and Filter	40
6.2.5. Manual	40
6.3. Setting up a New Mapping	41
6.3.1. Give Title and choose Target Schemas	41

6.3.2. Edit/view Info Tab	41
6.3.3. Add Generic Mapping Metadata	42
6.3.4. Adding Source Schema	42
6.3.5. Adding Sample Data	43
6.3.6. Adding Target Schemas	44
6.4. Matching Table Operation	45
6.4.1. Accessing the Matching Table	45
6.4.2. Adding a New Map	45
6.4.3. Adding a New Link	46
6.4.4. Copy and Deleting Maps and Links	47
6.4.5. How to comment	48
6.4.6. View Controls	48
6.5. Mapping Patterns	49
6.5.1. X3ML Constructs	49
6.5.2. Mapping Source Root to Target Domain	49
6.5.3. Simple Field Mapping (One to One)	50
6.5.4. Mapping to Paths: Introducing intermediate nodes	51
6.5.5. Mapping to Paths: Introducing constant expressions	52
6.5.6. Using variables	53
6.5.7. Joining Maps	53
6.5.8. Multiple instantiation	54
6.5.9. Mapping under conditions - If rules	55
6.6. Instance and label Generators	55
6.6.1. Why Instance and Label Generators, what do they do?	55
6.6.2. Defining instance generation functions/patterns (offline)	56
6.6.3. Adding generator file (info tab)	56
6.6.4. Opening up the generator editor	57
6.6.5. Specifying Instance Generators	57
6.6.6. Testing Transforms	58
6.6.7. Visualize transformed records	58
6.6.8. RDF visualizer	58
7. Data Mapping	59
7.1. Getty AAT	59
7.2. Period0	62
7.3. WGS84	64
8. Aggregation Pipeline Summary	66
9. Appendix 1: Properties of AO_Data_Resource	67

1. Introduction

The purpose of this User Guide is to provide a short introduction to the ARIADNEplus data aggregation pipeline. It defines, for the archaeological data providers, the process by which their data should be uploaded to the ARIADNE Content Cloud (AC for short), so that it appears in the ARIADNEplus Catalogue, and can be searched via the ARIADNEplus Portal. It provides a basis for Task 5.1 – Definition of the update procedure of the ARIADNEplus Data Infrastructure.

1.1. Related documents

- The specification of the AO-Cat Ontology, the living document where the ontology of the ARIADNE Catalogue is described in detail, available on the project infrastructure at <https://ariadne.d4science.org/group/ariadneplus-gateway/workspace?itemid=75a774d6-da40-43d4-96f3-184bca3c297e&operation=gotofolder>
- Data resource template – the properties defined in AO-Cat for an AO_Data_Resource (See also Appendix 1) <https://docs.google.com/spreadsheets/d/1Yqv1glbaHI8rNnhF9AY2zBjcTUSaHWxEwnSI6oaM6HM/edit?usp=sharing>

1.2. Structure of the document

The document is structured as follows.

- Section 2 outlines the purpose of the aggregation
- Section 3 introduces the AO-Cat
- Section 4 introduces the Activity Dash workflow dashboard
- Section 5 presents the options for how partners should supply their data to CNR-ISTI, and considers the frequency of supply
- Section 6 presents the 3M Mapping tool, which helps partners align their data to the AO-Cat (The data model used in the AC), and thereby to the CIDOC CRM
- Section 7 defines how data should also be mapped to standardised terms, to allow cross-search of the AC according to What, When and Where criteria
- Section 8 summarises the pipeline

2. Background: Why are we doing this?

This Section briefly outlines the requirements that are expected to be fulfilled by the ARIADNEplus Catalogue and more generally by the ARIADNE Content Cloud (AC for short) which the Catalogue is a part of.

There are two basic requirements:

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 ARIADNEplus Catalogue should support:

- *What* searches, searches based on a topic. For topics 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 have been defined by a User Needs study undertaken by the project (Deliverable 2.1), and will be implemented via VREs in d4Science.

Concerning enabling 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, categorized following standard vocabularies and properly connected to the relevant archaeological resources, as schematized in Figures 1 and 2, where each event category is annotated with the number of the sub-task of ARIADNEplus Task 4.4.0 addressing it:

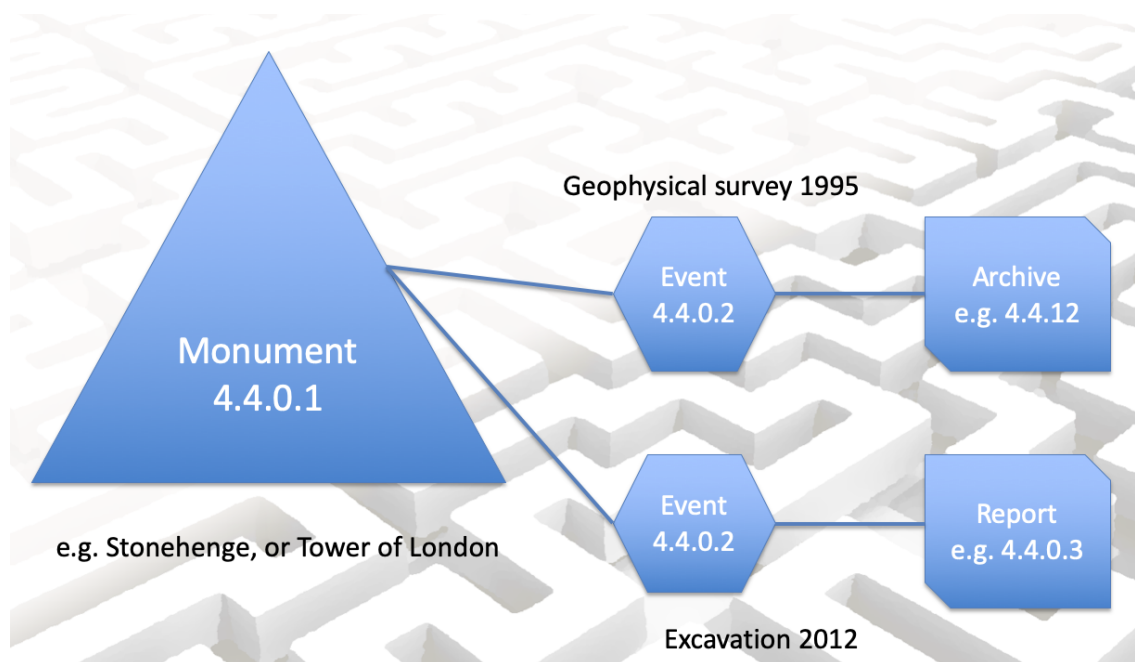


Figure 1. Entities and relationships in the archaeological domain

The diagram follows the UK MIDAS heritage standard terminology¹, 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 (Figure 2).

¹ <http://www.heritage-standards.org.uk/midas-heritage/>

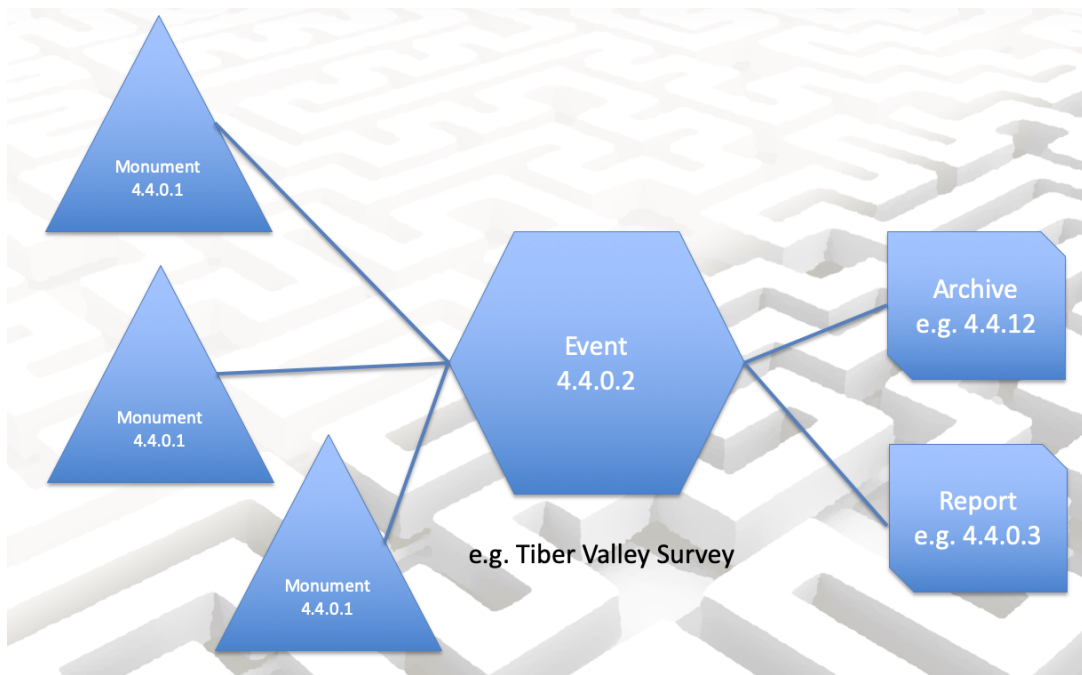


Figure 2. Entities and relationships in the archaeological domain - example of where a single recording event leads to the discovery of several new monuments

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. Matching the requirements: introduction to the AO-Cat

This Section provides a simplified explanation of how the requirements presented in Section 2 have been taken into account to define the AO-Cat. For further information see the AO-Cat Ontology.

All entities are identified by an IRI (Internationalized Resource Identifier, the identifier of resources on the web, previously known as URI, Uniform Resource Identifier) in the ARIADNE namespace.

The most general class of AO-Cat is AO_Entity. As such, AO_Entity has as instances all resources that have any role in the ARIADNE infrastructure. 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.
This will be displayed prominently in the portal, and so should be meaningful, such as the name of a site, or class of an artefact
- has_description, a textual description of the resource in natural language.
This will also be displayed prominently, and should provide as much information as possible; through the mapping it may be a merger of multiple fields

In addition, 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.

The AO-Cat is designed to catalogue several types of digital infrastructural resources. The most general infrastructural resource class is AO_Resource, representing all the digital resources the ARIADNE research infrastructure deals with. Any ARIADNE Resource has a type which can be specified in two different ways.

- One way is via the property has_type, which associates an instance of AO_Entity with a type represented as an AO_Concept.
- The other way is via the sub-classes of AO_Resource, also introduced next.

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 first acquired.
This might be the date the record was created in the native provider database
- was_modified, the date when the record of the resource was last modified.

This will be the date the record was last updated

- has_publisher, the agent responsible for making the resource publicly accessible.

This is usually the name of the organisation providing the data to ARIADNE, e.g. 'Archaeology Data Service'. It will be displayed as the "publisher filter" in the portal. It is therefore important that data providers choose the name by which they wish to be known, whether an acronym, or full name, and use it consistently, both for item level and collection level records.

- has_contributor, a contributor of a description of the resource to the AC.
This is usually the original contributor of the dataset to the partner data provider, either an organisation or an individual contributor. Thus for ADS, it could be 'Historic England' for example.
- has_creator, a creator of the resource.
For example, the person who created the dataset
- has_owner, the owner of the resource.
This is the organisation that holds legal ownership of the dataset
- has_responsible, any person who is scientifically responsible for the resource.
This may be the same as the person who created the dataset

These properties have been chosen based on the experience gained in the first ARIADNE project.

AO-Cat provides two subclasses of class AO_Resource to represent the main types of digital resources: AO_Data_Resource for data resources and AO_Service for services. In this document we are simply concerned with Data Resources (AO_Data_Resource).

Data resources are also characterized 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 that the resource refers to, 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 an entity that the data resource documents, including of course events. Notice that the documenting data resource may be an external resource, identified by a IRI, or DOI, or any other identifier, to capture the requirement that the ARIADNE AC should have links to distributed digital and paper resources in the outer world.

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

- `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 (see [Section 3.4](#)).

The taxonomy of aboutness properties in AO-Cat is given in Figure 3.

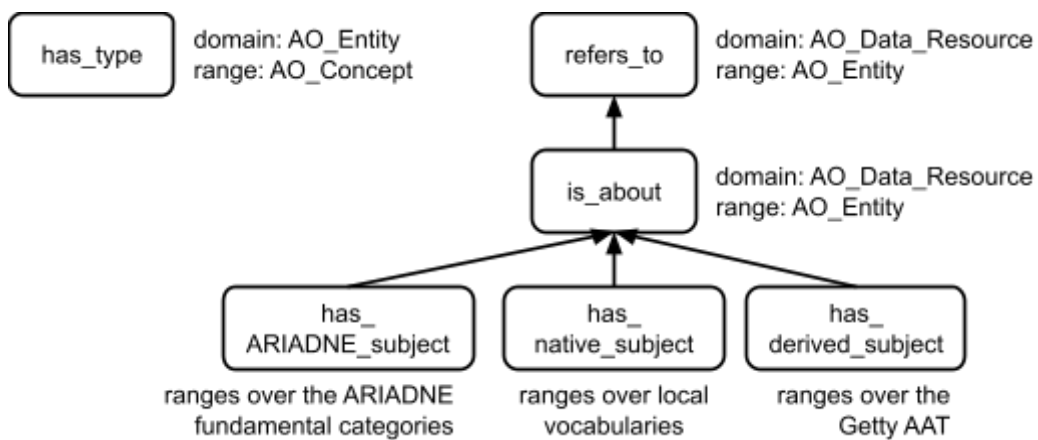


Figure 3.

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.

See Sections 3.2 and 3.3 below for more information on these important properties.

Several descriptive properties are also defined for data resources:

- `has_language`, the language of the resource;
should follow standard ISO abbreviated codes, such as 'en'

- was_created_on, the date in 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 kinds (the complete resource class taxonomy is given in Figure 4):

- 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.

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) it primarily carries information about. 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 ARIADNE needs to deal with, such as for instance the grey literature, published journal articles, radiocarbon dating resource, and possibly others, whose content may possibly be part of the AC.

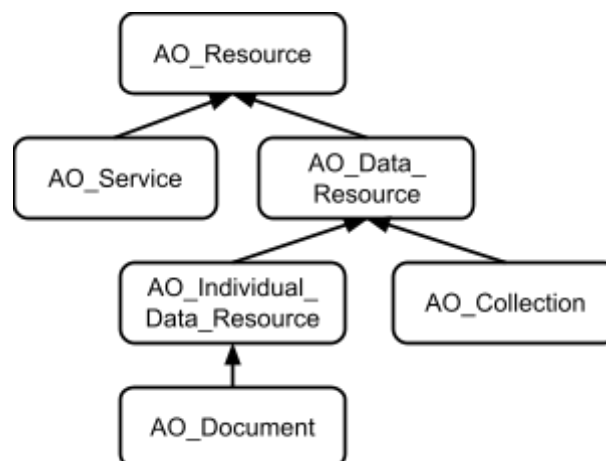


Figure 4.

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

3.1. 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 contextualizing 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 in the ARIADNE AC all physical objects that are relevant to ARIADNE. Three properties are defined for it:

- `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.

3.2. Defining Space

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

- `has_coordinate_system`
- `Has_place_name`
- `has_country_code`

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

- `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. No commitment is made by AO-Cat on a specific notation 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

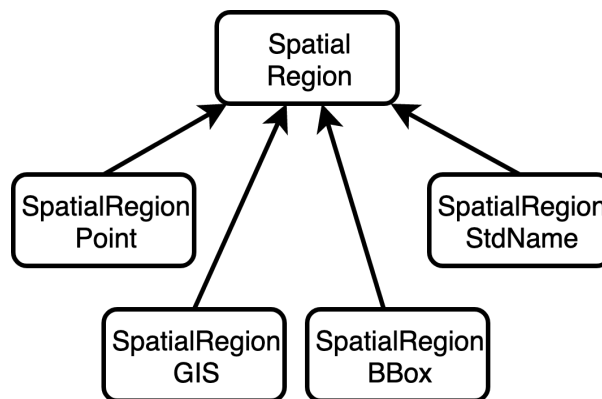


Figure 5

3.2.1. Defining spatial precision

It is useful to specify the precision of the spatial information provided, either because:

- The data provider has sensitive spatial data which they do not wish to make publicly available, maybe to protect the precise location of a site. Therefore they should only supply to ARIADNE the data at a resolution which they are willing to make publicly available. e.g. for the UK Portable Antiquities Scheme this is the parish administrative unit, and Geonames is then used to provide a centroid point for the parish, or
- The data provider may only hold data for site location at an imprecise level, so we need to make it clear when we plot locations on a map, that the point represents the site location to the nearest e.g. 10km.

In order to cope with these and with similar cases, we introduce a property:

- `has_spatial_precision`

associating a data resource with a specification of the precision that characterizes the spatial coverage of the resource. This precision is given as an instance of class `AO_Dimension` and should be specified as a numeric value e.g. 1km, 100km. Therefore `AO_Dimension` provides the two properties:

- `has_value`
- `has_unit`

both having a literal as value.

3.3. Defining time

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 representation of time intervals and their names, AO-Cat relies on the PeriodO gazetteer service² which allows to define a temporal interval as a web resource, associated to a label and to two absolute dates giving the earliest start and the latest stop of the interval. The service also allows to cluster period resources in collections, thus facilitating the exploration of the gazetteer data space³.

The AO-Cat class `AO_Temporal_Region` has time intervals as instances. Each such instance can be described by means of 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`.

² <https://perio.do/en/>

³ including about six thousand period definitions as of June 14, 2019.

For more information on mapping to space and time see [Section 7](#).

3.4. The ARIADNE fundamental categories - archaeological data types

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 ATT Thesaurus
- any other term used in the data of a provider.

In the ARIADNE portal the ARIADNE fundamental categories are defined as a very high level classification and used as a mechanism to filter searches according to types of data sets. The current ARIADNEplus fundamental categories - ARIADNE_subject - (and corresponding A+ sub-tasks) are:

- Site/monument (4.4.0.1)
- Fieldwork (Sub-task 4.4.0.2)
- Fieldwork report (sub-task 4.4.0.3)
- Scientific analysis (Sub-task 4.4.4, plus 4.4.1, 4.4.2, & 4.4.3)
- Date (4.4.5)
- Artefact (4.4.7)
- Coin (sub-set of 4.4.7)
- Building survey (4.4.9)
- Maritime (4.4.11)
- Fieldwork archive (4.4.12)
- Inscription (4.4.13)
- Rock Art (sub-set of 4.4.13)
- Burial (4.4.14)

These categories emerged as *ad hoc* pragmatically defined data clusters in the first ARIADNE project and they have evolved and increased as the WP4 subtasks gather information about data categories commonly held by A+ partners. Although the AO-Cat allows the specification of several ARIADNE_subjects for a resource we are encouraging partners to choose the most appropriate (at the most detailed or granular level available) to maximize the value of this category for filtering purposes in the portal.

3.5. Events and activities

Much of the archaeological research activity concerns fieldwork. Fieldwork results in digital data resources being collected or created for the purpose of documenting the research activity, presenting its results, or arguing in various ways, for instance expressing hypotheses or refuting other scholars' hypotheses. Digital data resources may also be the result of analyses that are conducted in a research laboratory.

In AO-Cat, these research activities are generally modelled as activities, in the sense of the “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].

4. Using the Activity Dash tool

Activity Dash is an online web application, accessed through any modern web browser, aiming in organizing and tracking workflows with respect to their several processes (activities). **As soon as you have authorisation to proceed with aggregation of your data a dedicated workflow will be set up by the sub-task leader** (generally Julian Richards). End users of the system are capable of observing the progress developed under one or more workflows and updating it, when required. In general, the application combines:

- i) A workflow manager and tracker;
- ii) A collaborative environment;
- iii) A notification means for stakeholders;
- iv) A reference point for stakeholders, to watch the progress done on workflows.

All workflows within the system are grouped by projects, where access is regulated by user groups, defined by some project owner user. Following the same pattern, each workflow within a project has its own user group defined by the workflow owner. The workspace environment offered by the system, presents the workflow as divided in activities, which are further divided in tasks. The workflow owner, assigns users to activities, and each activity assignee, assigns other users to tasks. All stakeholder users can simultaneously edit or extend the workflow, without worrying about missing potential changes being made by others on their shared workspace. This is mainly achieved, due to systems capability of syncing the shared workspace view, among all users behind different browsers.

Activities within the workflow are capable of holding associations between others. This is mainly done when some activity depends on the completeness of other activities. The same can also be applied for tasks, always within the same activity. The system is capable of automatically informing stakeholders, under specific circumstances, about the completeness of some activity and any new activity or task assignment or dismissal.

4.1. Getting Started

The necessary conditions for any user to access Activity Dash, is to be registered in D4Science and have access to the AriadnePlus Mappings VRE. As soon as these two criteria are met, the system can be reached by visiting the link https://ariadne.d4science.org/group/ariadneplus_mappings/activity-dash.

Please note that every member of the AriadnePlus Mappings VRE should access Activity Dash at least one time. In order for Activity Dash to perform properly, all users that have been used in the constructed workflows should have their email attached to their profile. Activity Dash will automatically retrieve the necessary profile information from D4science at the very first time that each user accesses Activity Dash. Even though most of the required information of the AriadnePlus Mappings VRE members have massively been imported to Activity Dash offline, their emails

were not possible to be included, which means that these users should access Activity Dash at least once.

There are two different modes to choose, “Embedded” mode and “Full Screen” mode. The differences between these two modes are visual and practical. “Embedded” mode will display Activity Dash wrapped within the workspace of D4Science, while “Full Screen” mode will be displayed in an individual browser tab. The “Full Screen” mode takes advantage of all the available space, due to the lack of the D4Science workspace wrapper.

4.2. Accessing Project

The first view of the application, simply displays the available projects to visit. As mentioned before, workflows are grouped by projects. To get inside a project, first click on it in order to select it and then click on the “Open” button, located at the bottom left corner. Please note that the button will appear to be grey, if the project has not been selected first.



The screenshot displays the Activity Dash interface. At the top, there is a teal header with a menu icon, the 'FI' logo, the user name 'Vangelis Kritsotakis', and a profile picture. Below the header, the main content area is titled 'Projects' and includes a 'Show All Projects' toggle switch. The featured project is 'ARIADNEplus', which includes a profile picture of Julian Richards and a detailed description of the project's scope and goals. A teal 'OPEN' button is positioned at the bottom left of the project card. The footer contains the text 'Terms of Use | Privacy Policy | © 2019-2020 FORTH-ICS | Licensed under the EUPL' and logos for FORTH, IS, and the Center of Cultural Informatics, along with the text 'Designed and developed by CCI'.

4.3. Accessing Workflow

The new view displays all the workflows within the project. Searching capabilities are available through the UI components located at the top right corner just above the data table. The buttons at the bottom left corner are activated as soon as a workflow is selected. These buttons can be used for opening, cloning or exporting (to file) the workflow that has been selected. The picture below explains the functionality described here for the current view.

The screenshot shows the 'ARIADNEplus - Workflows' page. The interface includes a header with a menu icon, the 'FI' logo, and the user name 'Vangelis Kritsotakis'. Below the header, there is a navigation bar with a back arrow, the page title, a 'Show Ongoing Only' toggle, a search bar, and a 'Select Column Filter' dropdown. The main content is a table with columns: Name, Description, Completeness, Owner, and Avatar. The table lists several workflows, with 'AIAC FASTI collection - LIVE' highlighted. Below the table, there are 'Rows per page' and '1-10 of 20' indicators. At the bottom, there are three buttons: 'OPEN', 'CLONE', and 'EXPORT TO FILE'. The footer contains 'Terms of Use | Privacy Policy | © 2019-2020 FORTH-ICS | Licensed under the EUPL' and the 'FORTH ICS' logo.

Annotations in the image explain the following features:

- Sorting by column:** Points to the column headers in the table.
- Filter that hides all the completed workflows:** Points to the 'Show Ongoing Only' toggle.
- Search by text:** A dashed box containing two sub-annotations:
 - Type some text to search:** Points to the search input field.
 - Select what to search for (name, description, completeness, and owner):** Points to the 'Select Column Filter' dropdown.
- Opening Selected Workflow:** Points to the 'OPEN' button.
- Cloning Selected workflow:** Points to the 'CLONE' button.
- Exporting selected workflow to file:** Points to the 'EXPORT TO FILE' button.
- Select the number of workflows to show per page:** Points to the 'Rows per page' dropdown.
- Next / Previous Page:** Points to the pagination arrows.

Any workflow is opened by selecting it and then clicking on the “Open” button (bottom left corner). When this is done, the workflow is loaded on a new view, where users can observe or edit the progress or information details of the workflow’s components.

4.4. Explaining the Workflow View

The workflow view shows the workflow itself with respect to its activities (processes constituting it) and tasks (sub-processes constituting the activities). Please note that tasks are optional units, which means that an activity might or might not be divided into tasks. This view shows all the details of the workflow, including dependencies and conversations, which will be discussed later on in this chapter.

The screenshot shows a workflow management interface with the following callouts and their corresponding UI elements:

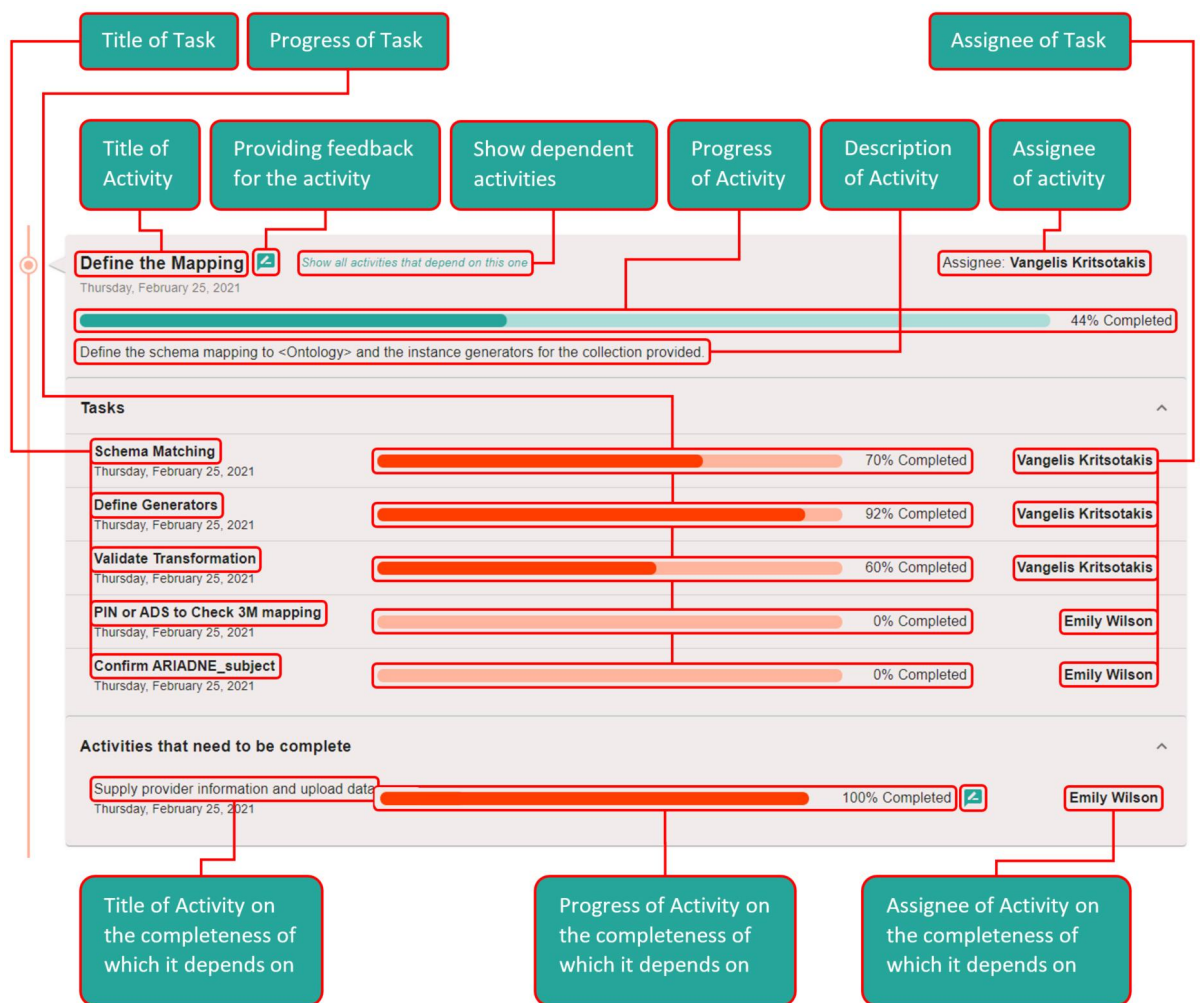
- Activities constituting the workflow:** A legend on the left side of the page showing three activity states: Completed (green square), Ongoing (orange square), and Without any progress (grey square).
- Breadcrumb to go back to the list of workflows or projects:** The breadcrumb path at the top: "Projects > Demonstration - Workflows / Ariadne Plus Demo Workflow".
- Providing feedback for the workflow / activity:** A feedback icon (a speech bubble with a checkmark) located next to the activity name.
- Online users currently editing the workflow:** A user avatars indicator showing two users and a "+2" sign.
- Switching to "Edit" mode:** An "Edit Workflow" button located at the top right of the workflow list.

The main content area displays four activity panels, each with a title, description, progress bar, assignee, and expandable sections for tasks and dependent activities:

- Supply provider information and upload data:** Assignee: Emily Wilson, 100% Completed.
- Define the Mapping:** Assignee: Vangelis Kritsotakis, 44% Completed.
- Convert data to RDF:** Assignee: Molly Brown, 0% Completed.
- Public Release:** Assignee: Mark Williams, 0% Completed.

At the bottom of the page, there is a footer with the following text: "Terms of Use | Privacy Policy | © 2019-2020 FORTH-ICS | Licensed under the EUPL". Logos for FORTH, ICS, and the Center of Informatics and Telematics are also present, along with the text "Designed and developed by CCI".

Each activity panel shows information about the activity itself (name, description, progress and assignee), the tasks constituting it (if there are any) and the activities, on the completeness of which it depends on. The tasks and the activities on which it depends on, are listed in separated collapsible panels. By clicking on them they are expanded revealing the respective information. The picture below explains each part of the activity panel.



4.5. Applying Changes on the Workflow

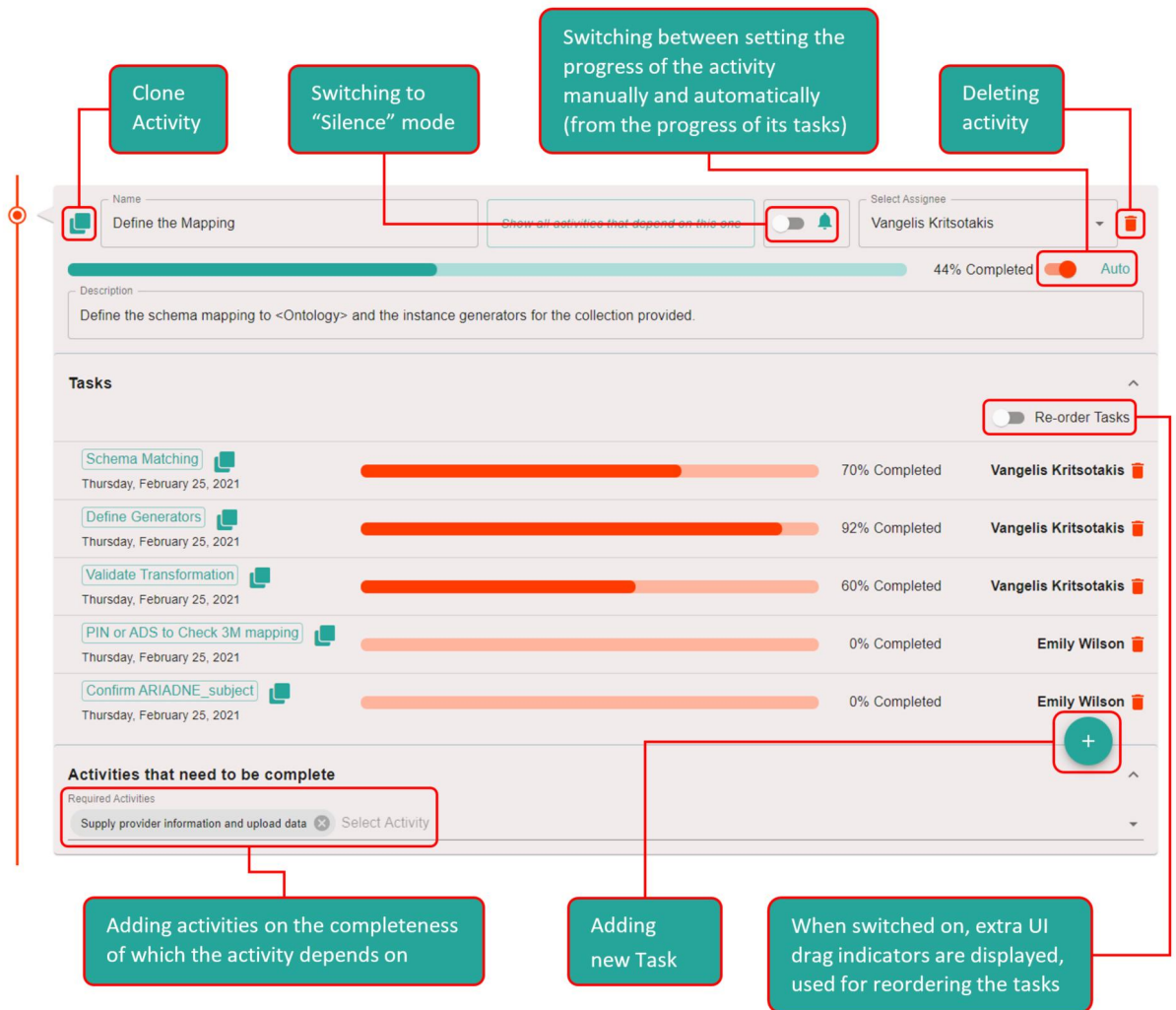
Each user can only apply changes on the activities or tasks where he/she is assigned to. The activity assignee is able to edit the whole activity (including its tasks), and the task assignee is only able to edit that task. By clicking on the switch located at the top right corner of this page, the mode changes from “View” to “Edit”.

When in edit mode the user can apply several actions:

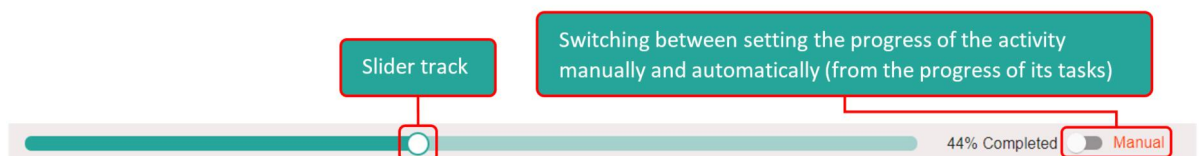
- edit activity or a task’s information (including the progress done);
- clone an activity or task;
- delete an activity or task;
- add new task into an activity;
- switch to silence mode, which will make the system to cease automatically notifying users about new assignments or progress changes (that last one is useful when using an activity as template and cloning it many times).

Furthermore, if the activity consists of many tasks, then it is possible to tell the system to always calculate activity’s progress from the progress done on its tasks.

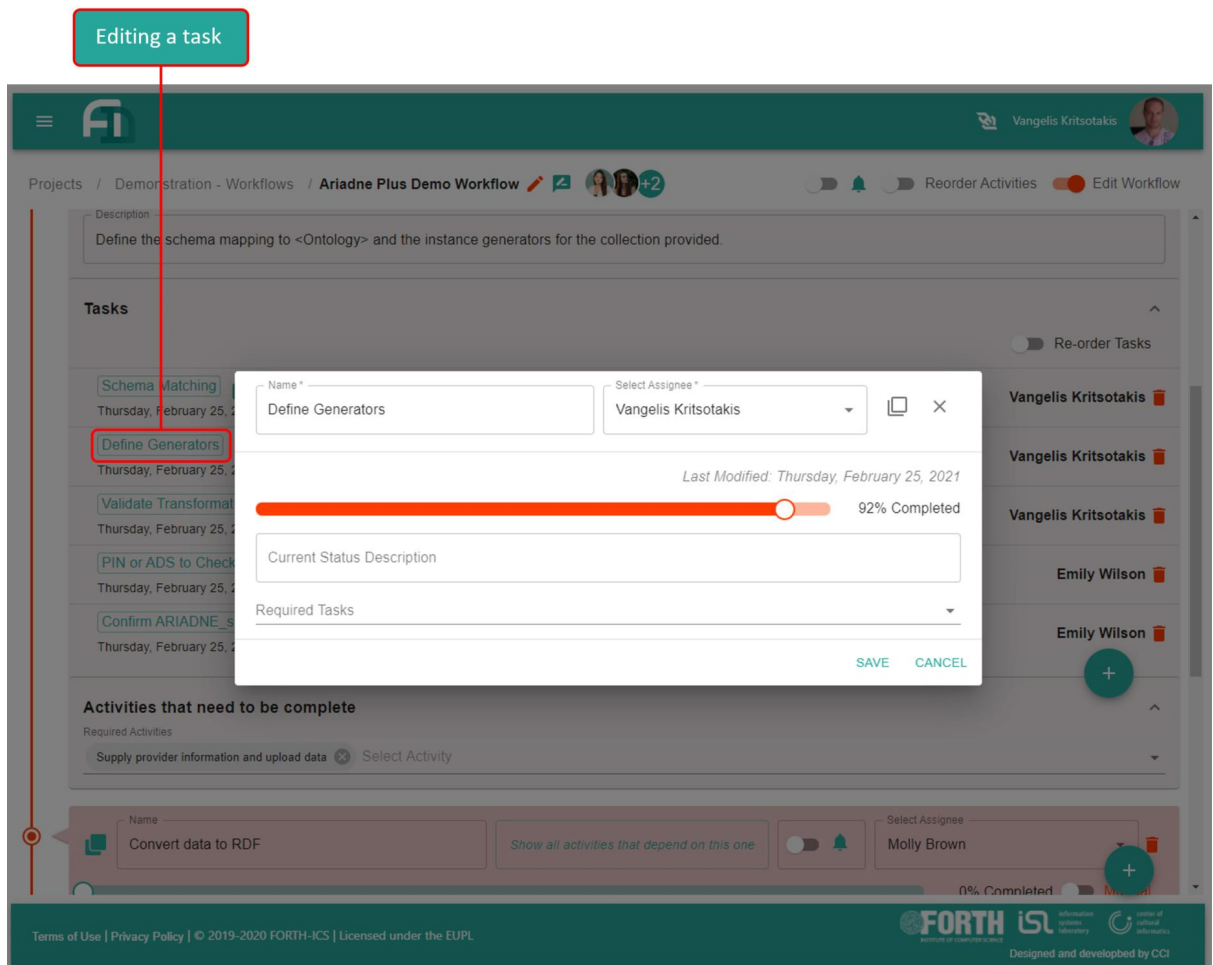
Below is a picture of an activity panel, when in edit mode, with basic explanation of the extra UI buttons.




The progress of the activity can directly be changed when the switch that enables the calculation of the progress, based on the respective activities, is off. This is done by dragging the slider track, on the progress bar, to the desired percentage value. If the switch mentioned above is enabled, then the activity's progress is calculated automatically by the progress of its tasks. In that case the only way of changing the progress, is to edit its tasks and change their progress one by one.

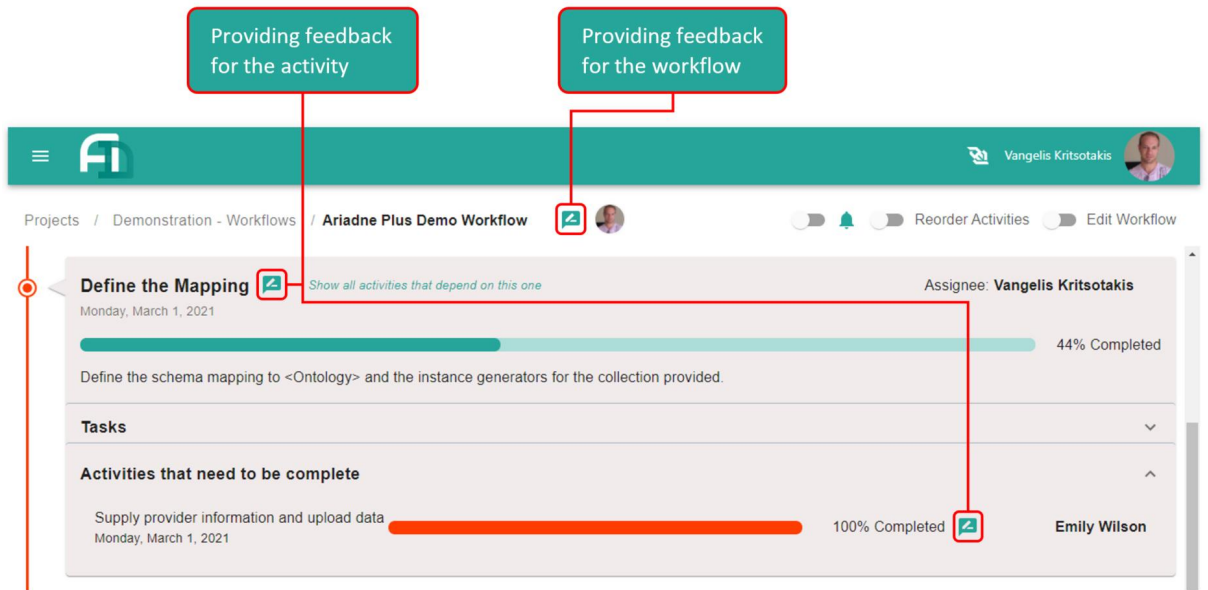


To edit a task (that includes changing its progress done), simply click on its name and edit the fields provided on the displayed dialog.

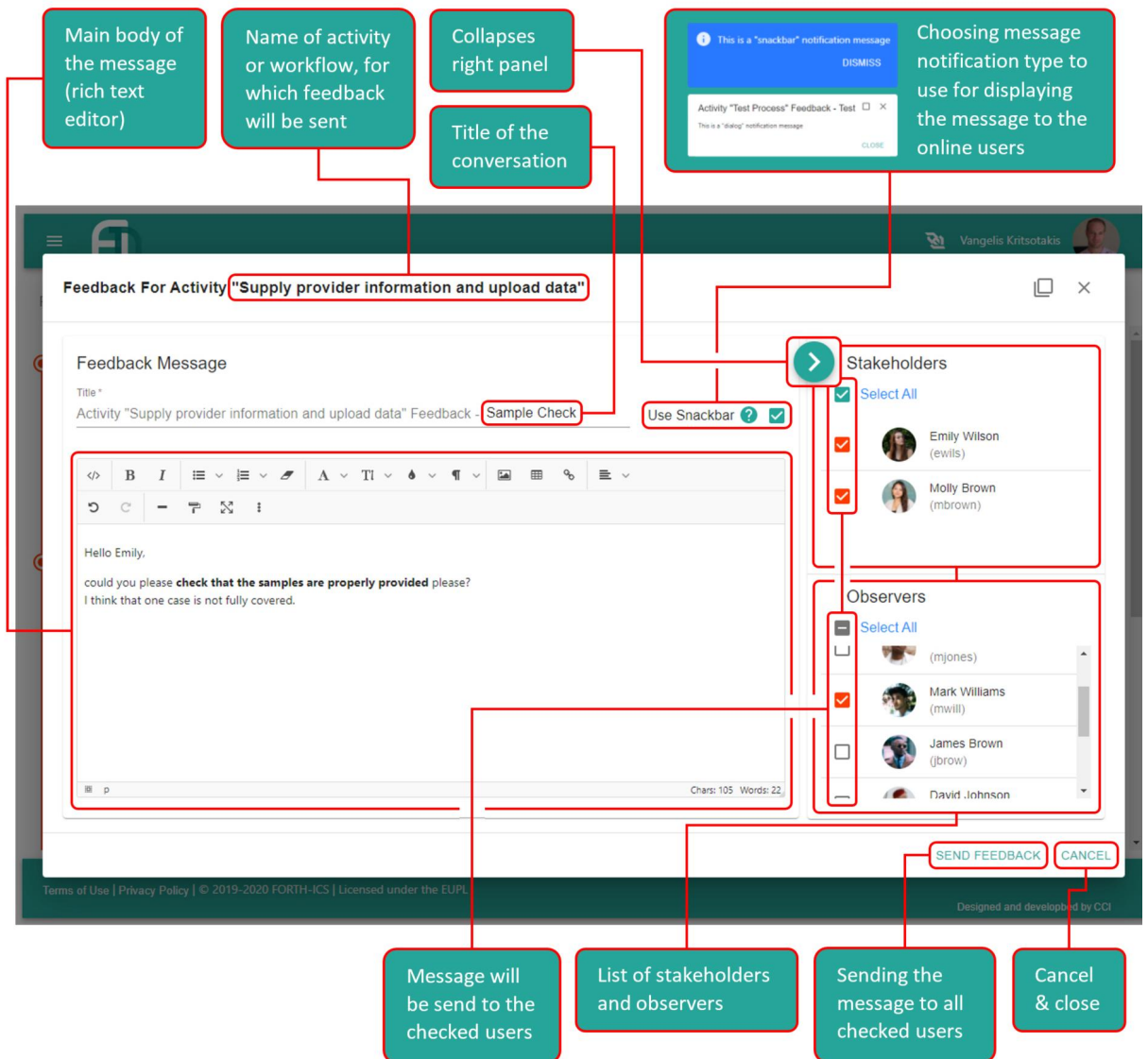


4.6. Providing Feedback

Each user that can access a workflow, is able to provide feedback regarding the workflow itself or any of its activities. Feedback in Activity Dash has the form of conversation (like chat), where many users can exchange information or discuss issues. These users can either be the stakeholders (in case of an activity feedback, those are the assignees of the activity and its tasks) and the observers (users that can watch the workflow but are not assigned to anything on it). Feedback is applied by clicking on the  icon, which is located next to the title of the activity or the workflow, for which feedback has to be provided.



As soon as the “Feedback” icon is clicked, a new dialog will be displayed on top of the UI, asking the user to enter a title for the new feedback conversation that is about to be constructed and the body of the initial message to be sent. On the right side of this dialog, there will be two lists of available users to be selected (checked). The first list holds the stakeholders and the second list holds the observers. As stakeholders, in case of activity feedback, are considered the assignees of the activity and its tasks and, in case of workflow feedback, are considered the users that are assigned some activity or task within that workflow. Finally, the observers are the rest of the workflow members (who are not assigned to anything). The message will be sent to all the selected users (no matter whether they are in the list of stakeholders or observers) by clicking on the “Send Feedback” message (please note that at least one user has to be selected in order for this button to be enabled).



As soon as the message is sent, a new conversation containing one message is created by the system. In case of activity feedback, a new collapsible panel is added in the respective activity's panel, which holds all the feedback conversations of this activity.

Projects / Demonstration - Workflows / Ariadne Plus Demo Workflow

Supply provider information and upload data Show all activities that depend on this one Assignee: Emily Wilson

Monday, March 1, 2021

100% Completed

Note when data provider information has been supplied and data uploaded

Tasks

Conversations

Activity "Supply provider information and upload data" Feedback - Sample Check

HRID: 199985

Mon Mar 01 2021 11:56 AM

Collapsible panel, holding all the conversation feedbacks associated to the activity

Any user that has access to the workflow can click on any conversation feedback and read it. When this is done, the perspective view will be changed.

Gets user back to the workflow

Name of activity or workflow (in case of workflow feedback)

Name of workflow

Conversation

Title: Activity "Supply provider information and upload data" Feedback - Sample Check

Workflow: Ariadne Plus Demo Workflow

Last updated on Mon Mar 01 2021 11:56 AM

Participants: Emily Wilson (ewils), Molly Brown (mbrown), Mark Williams (mwill), Vangelis Kritsotakis (vkrits)

Vangelis Kritsotakis

Mon Mar 01 2021 11:56 AM

Hello Emily.

could you please check that the samples are properly provided please?
I think that one case is not fully covered.

REPLY

GET ME TO ACTIVITY "SUPPLY PROVIDER INFORMATION AND UPLOAD DATA"

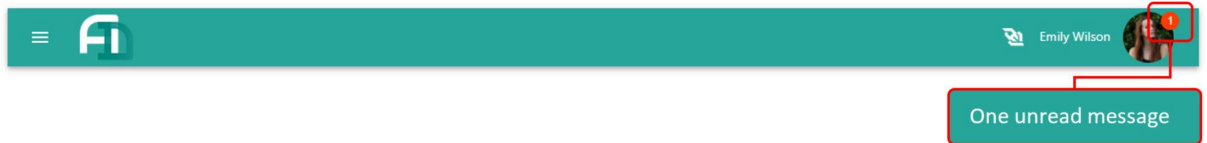
Displays the feedback dialog to reply to this

Gets user either to the workflow or to the activity (depends on the feedback type)

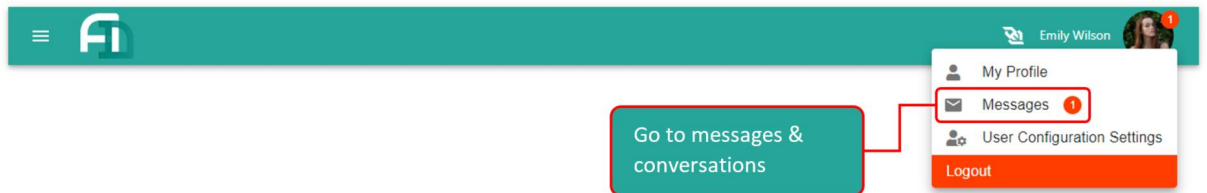
4.7. View and respond to messages or conversations

Users that have received a feedback message and thus are participating in the conversation, will get an email containing the message, a numerical badge icon notification, increased by one, on their avatar picture denoting the number of unread messages, and if online a notification message on the top right corner. The above will

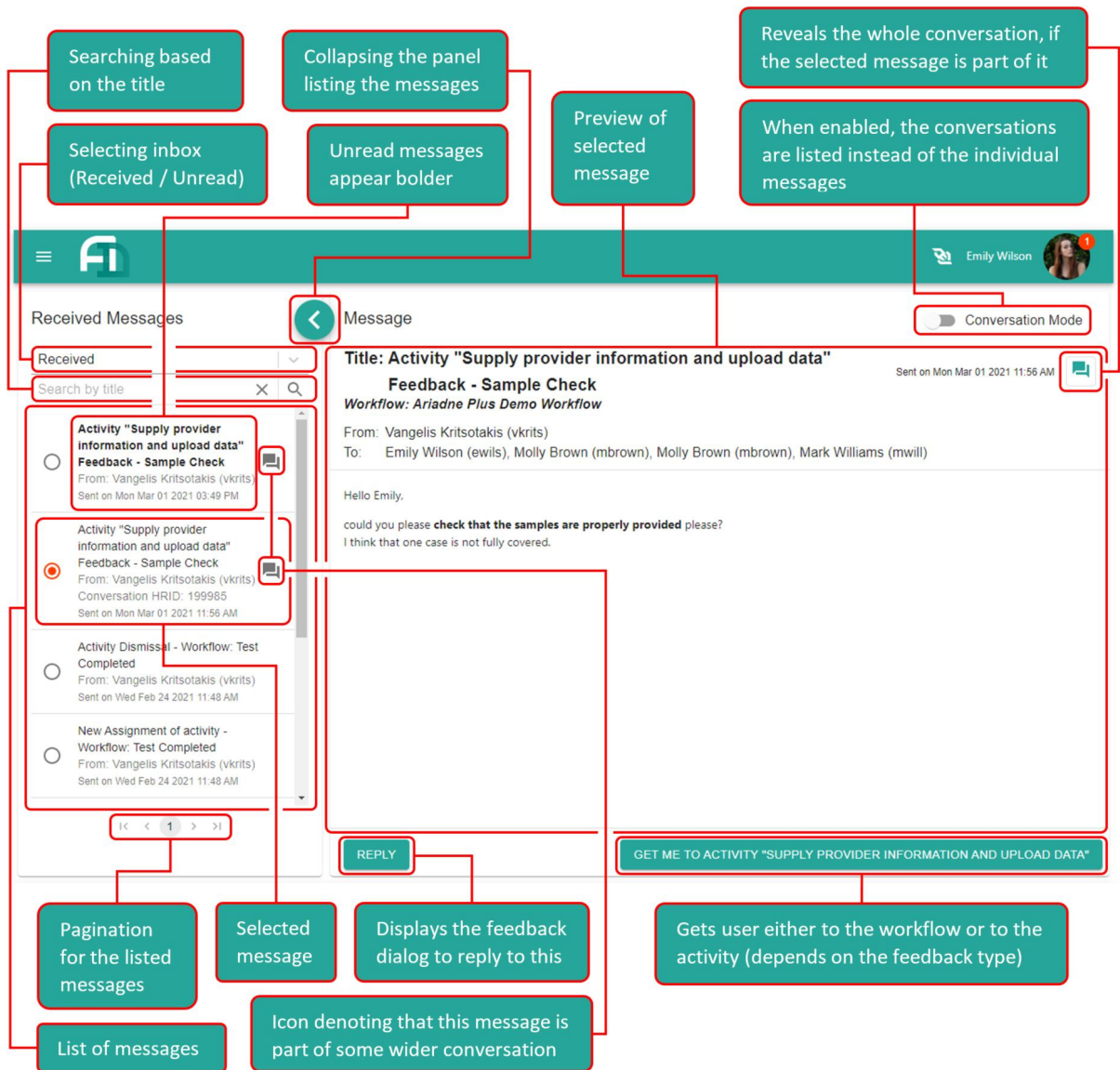
be applied with the default configuration settings of the user. User configuration settings can be changed by the users to meet their personal needs.



By clicking on the user's avatar image and selecting the “Messages” option, the system will display a new view used for searching or responding to messages or conversations.



The new perspective view initially lists all the individual messages received by the user, no matter whether they are sent by the system automatically or by some other user as part of a wider conversation (left panel). The messages are listed on the left panel while the selected message is previewed on the main panel. On the top right corner there is a switch, which switches between individual messages and conversations, and at the bottom of the main panel are located two buttons, one to reply if possible (you cannot reply to messages sent by the system) and another one which navigates to the respective activity or workflow, if that message is part of some conversation feedback.



The user can switch to “Conversation” mode view, to search and preview conversations. This view is similar to the “message” mode view described earlier, with only difference that the panel on the left lists conversations instead of individual messages, and when the user selects one, the whole conversation is previewed in the main panel. When some reply is received while the conversation is previewed, it is automatically updated with the new message.

The screenshot displays the Activity Dash interface. At the top, there is a teal header with the 'Fi' logo on the left and the user's name 'Emily Wilson' on the right. Below the header, the interface is split into two main sections.

List of conversations: On the left, a sidebar titled 'Participating Conversations' contains a search bar and a list of five conversation items. The first item, 'Activity "Supply provider information and upload data" Feedback - Sample Check', is selected and highlighted with a red box. The other items are 'Activity "First Case" Feedback - Too Many', 'Activity "First Case" Feedback - sss', 'Activity "First Case" Feedback - 33344444', and 'Activity "First Case" Feedback - Hello fellows'.

Preview of the conversation: The main area on the right shows a detailed view of the selected conversation. It has a title 'Activity "Supply provider information and upload data" Feedback - Sample Check' and a subtitle 'Workflow: Ariadne Plus Demo Workflow'. Below this, it lists participants: Emily Wilson (ewils), Molly Brown (mbrown), Mark Williams (mwill), and Vangelis Kritsotakis (vkrits). The conversation history shows three messages:

- Vangelis Kritsotakis (11:56 AM): Hello Emily, could you please **check that the samples are properly provided** please? I think that one case is not fully covered.
- Emily Wilson (03:46 PM): Hello, I have updated the samples. They should be OK now.
- Vangelis Kritsotakis (03:49 PM): Yes I can see they 're fine now. Thanks

Conversation mode is on: A red callout box points to a toggle switch labeled 'Conversation Mode' which is currently turned on. At the bottom of the conversation view, there are two buttons: 'REPLY' and 'GET ME TO ACTIVITY "SUPPLY PROVIDER INFORMATION AND UPLOAD DATA"'. A red box also highlights the 'Conversation' title and the 'Conversation Mode' toggle.

4.8. More Detailed Instructions

More instructions that cover further functionality of Activity Dash, like creating projects and workflows can be obtained from the user manual, a link of which can be found from within the Activity Dash, when opening the main left menu.

5. Supplying Data to CNR

As a content provider in ARIADNEplus, partners are expected to make their metadata available so that it can be collected, transformed into the ARIADNEplus data model (AO-Cat), stored on a triplestore based on the [GraphDB technology](#), and enriched with links to [PeriodO](#) and [Getty AAT](#).

This section describes what partners need to provide so that the ARIADNEplus aggregation system (managed by ISTI-CNR) can be configured to apply the workflow (collect, transform, enrich) to their records. It is strongly recommended that they also check the results before their content is made publicly available.

Figure 6 shows the data flow that will result in the availability of enriched, AO-Cat compliant metadata records in the ARIADNEplus portal.

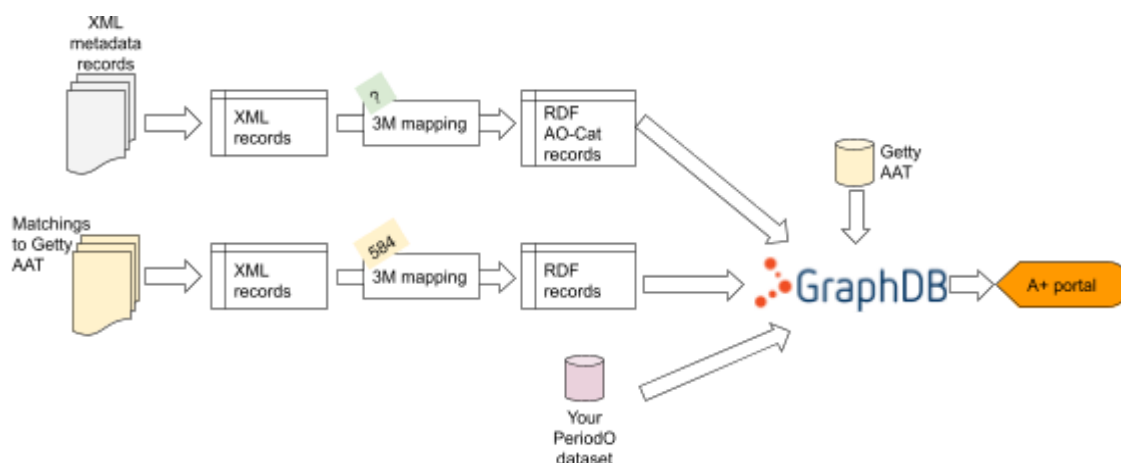


Figure 6. The aggregation workflow applied to each set of records

5.1. In a nutshell

- Get the go-ahead to proceed from the task leader who will set up a dedicated workflow in Activity Dash (Section 4).
- Make your sample XML records available (see section 5.2 “Metadata Ingestion”)
- **Complete the online questionnaire from CNR** which provides essential partner information for the aggregation process, following this link: <https://forms.gle/rhLWjVEiwb1AsV94A>
- Use Activity Dash to Inform the aggregation team if you cannot provide normalised dates: you will be given instructions on how to use a tool that normalises them and how provide the output of the tool along with your XML records (if you cannot modify the records themselves to include the generated normalised dates)

- Activity Dash should be used to inform the aggregation taskforce which is the mapping for your metadata sets and when it is ready (see section 5.3 “Metadata Transformation”)
- Activity Dash will be used to record when you are ready with your matchings to Getty AAT and your PeriodO dataset (see section 5.4 “Metadata Enrichment”)
- You’ll be notified via Activity Dash when the records are ready and you will be able to preview them in the “ghost”/ or staging portal, only accessible to the members of the consortium (see section 5.6 “Check your content!”)
- If everything is ok, use Activity Dash to provide partner “sign-off” to your records. This gives CNR the green light to publish them in the public ARIADNEplus portal.

5.2. Metadata ingestion

The ARIADNEplus aggregation system is based on the D-Net software toolkit (implemented and maintained by ISTI-CNR), which provides built-in features for collecting metadata records in different ways. The main options are⁴:

- **OAI-PMH:** OAI-PMH is a standard protocol for metadata exchange. The aggregator can collect the whole content or only from one or more specific OAI sets. Incremental collection of records is supported if your OAI-PMH publisher supports it.
- **SFTP:** there should be one XML file per resource (i.e. not one single, big file). Incremental collection is supported. You are responsible for the SFTP server. Public key or username/password authentication modes are supported.
- **FTP(S):** there should be one XML file per resource (i.e. not one single, big file). Incremental collection is supported. You are responsible for the FTP server.
- **Dumps:** you can upload the records in your folder [here on the d4science workspace](#) (Workspace > VRE Folders > ARIADNEplus_Mappings > Metadata Ingestion - login required). If you have problems in accessing the folder or login, please open a ticket to the [helpdesk](#).

IMPORTANT: XML files will be transformed with the 3M mapping you have defined. Please ensure the export format is the same as the one you used when defining the mapping. In the case of OAI-PMH, each oai:record will be passed to the transformation.

IMPORTANT: Each XML file should have one XML field containing a unique identifier. In the case of OAI-PMH, it can be the oai:identifier. The availability of a unique identifier per resource is also a requirement of the 3M mapping.

IMPORTANT: if you have several collections that must be transformed with different 3M mappings, then you should provide the relative records in different sets/folders/dumps.

⁴ If you are not sure what to choose or if none of the options suits, please open a ticket on the [helpdesk](#).

5.3. Metadata Transformation

The aggregator will automatically transform the collected records to the ARIADNEplus data model by applying the 3M mapping you have defined with the [X3ML mapping tool](#) (see Section 6).

If your XML records already contain normalised dates, then your 3M mapping must properly map the values into AO-Cat temporal coverages.

If instead your dates are not normalised, ARIADNEplus is offering a tool, called [Timespans](#), that automatically performs normalization based on a wide set of date patterns commonly used in archeology. The tool will produce as output a modified version of your XML records, including the normalised dates. You might need to update your 3M mapping to properly map the normalised dates into AO-Cat temporal coverages (“from” and “until” properties).

5.4. Metadata Enrichment

You have created the matching between your local terms and Getty AAT with the [Vocabulary Matching Tool](#).

Save the output on your folder [here on the d4science workspace](#) (Workspace > VRE Folders > ARIADNEplus_Mappings > Matched Vocabularies - login required). If you have problems in accessing the folder or login, please open a ticket to the [helpdesk](#).

You can also have multiple mapping files, if it is easier for you to manage. [See what UoY-ADS has done in their folder](#).

Please note that the workspace supports versioning: if you have an updated version of the mapping, upload it with the same name of the previous version. The old version will still be accessible by clicking on the file with the right button of the mouse and then select “Versions”.

The files will be collected by the ARIADNEplus aggregator, converted into RDF by applying the 3M mapping 584.

See also [Section 7.1](#).

To enrich your collection with the PeriodO URIs (e.g. fill the values in aocat:has_period), you just need to communicate the URL to your PeriodO dataset.

See also [Section 7.2](#).

IMPORTANT: if you are already using Getty AAT and PeriodO URIs in your records, then you can directly map them into has_derived_subject and aocat:has_period, respectively, with your 3M mapping.

5.5. Check your content!

When the process is completed, you'll be asked to access the [ARIADNEplus "ghost" \(staging\) portal](#), where you will be able to preview how the records will look in the ARIADNEplus portal.

If you are an RDF expert, we can also give you access to the [ghost GraphDB](#), so that you can explore the RDF graphs and run the SPARQL queries you prefer. To request access to the ghost GraphDB, please open a ticket on the [helpdesk](#) and specify your IP address (or address range).

We will publish your records on the GraphDB serving the public SPARQL endpoint and the public [ARIADNEplus portal](#) only when you give us your green light!

5.6. Frequency of updates

In the initial phase, we will perform several rounds of collection, transformation, and enrichment in order to tune all the mappings. Once we reach stability, we can agree on the frequency of execution. The ARIADNEplus aggregator can be configured to aggregate the records automatically according to a given schedule: once a week, once a month, or whatever fits best your use case. Examples of options might include:

- I have an OAI-PMH publisher that supports incremental harvesting and deleted records. Records are inserted/deleted/updated daily. I would like to see changes in the ARIADNEplus portal after one week → let's schedule an automatic incremental aggregation once a week
- I have an FTP server where records are automatically published once every 6 months (e.g. January and June), existing records are overwritten every time, records to be removed are deleted from the FTP server → let's schedule an automatic complete re-aggregation every 6 months (e.g. February and July).
- I manually upload the records on the d4science workspace when I know there are relevant records for ARIADNEplus → No schedule: inform WP4 leader of the new records when you are ready.

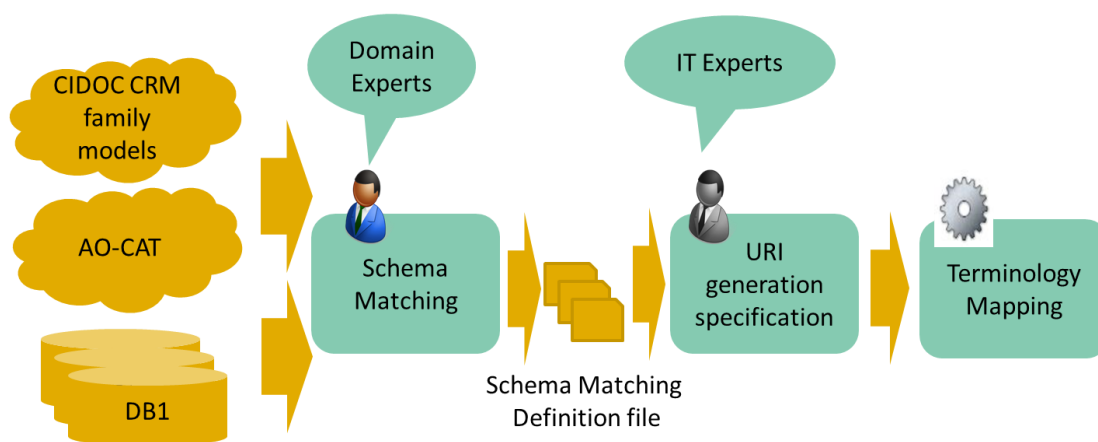
6. Using the 3M Mapping tool

The X3ML Toolkit is a set of small, open source, microservices that follow the SYNERGY Reference Model⁵ of data provision and aggregation. They are designed with open interfaces and they can be easily customized and adapted to complex environments. The key components of the toolkit are:

- Mapping Memory Manager - 3M is a tool for managing mapping definition files. It provides a number of administrative actions that assist the data supplying partners manage their mapping definition files.
- 3M Editor - is the interface tool that allows the domain experts to build mappings. It provides:
 - o Source and target agnostic mapping facility
 - o Guided mapping according to deployed ontology's logic
 - o Comment and justification facility
 - o Mapping storage
 - o Separated instance generation practice for IT professionals
- X3ML Engine - The X3ML Engine realizes 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 is responsible for transforming the source document into a valid RDF document which corresponds to the input XML file, with respect to the given mappings and policy.
- RDF visualizer – The visualizer allows a fast way to inspect the transformed records.

The X3ML Toolkit is deployed for ARIADNEplus on d4science at:

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



⁵ <http://www.cidoc-crm.org/Resources/the-synergy-reference-model-of-data-provision-and-aggregation>

6.1. Quick Start Guide

1. Start a new mapping: Copy **Mapping 594** which acts as a template and has the basic necessary files already loaded.
2. Open the newly created mapping file and in the **Info tab** select EDIT:
 - a. Change the Title of the mapping. It is suggested that you use the provider name in the mapping title.
 - b. In the Sample data and Generator policy section, Source record, load a file with a few sample records.
 - c. Fill in as much metadata as possible.
3. **Matching Table Tab**: Proceed to the matching of every source node to the respective AO-CAT node. Details on the matching are available in Section 5.4.
4. **Generators Tab**: Define the Instance and Label Generation functions. The manual with details on the existing functions is available at: https://mapping.d4science.org/3M/Manuals/en/X3ML_Generators_Manual.pdf
It is mandatory to give ONE instance generator to each class and ONE label generator. Label generators can be more than one.
5. **Transformation Tab**: Run the transformation on the sample records.
6. Check the console for errors and go back to steps 3 and 4 to fix them.
7. View transformed records in order to ensure that the transformation did what you had in your mind.

6.2. Basic Interface Operation

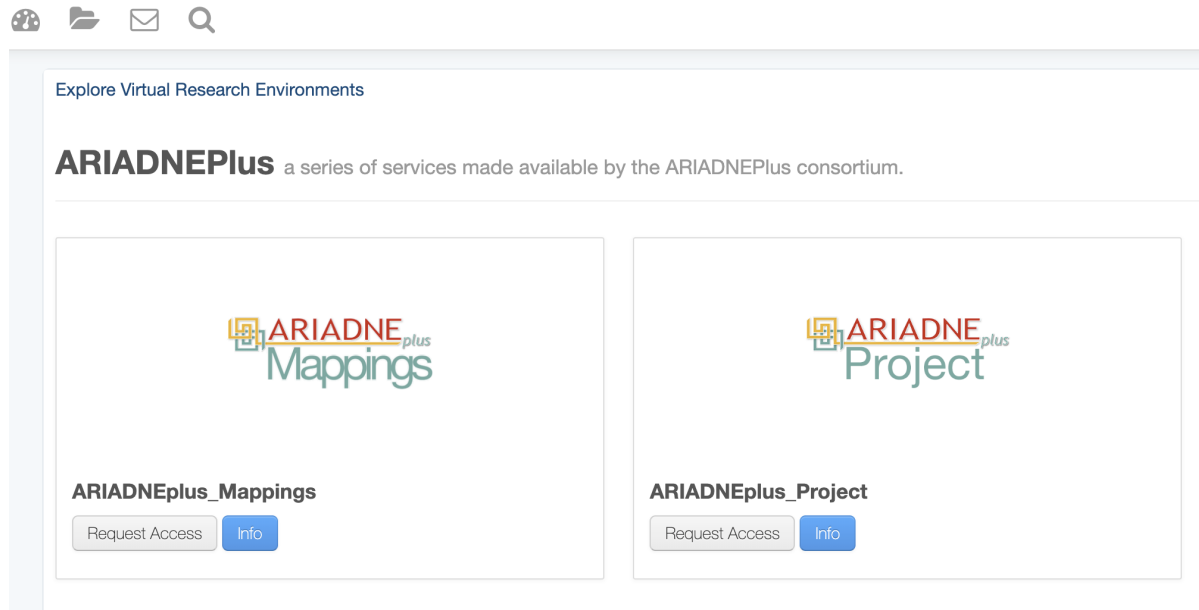
6.2.1. Login

Login using your d4science credentials at:

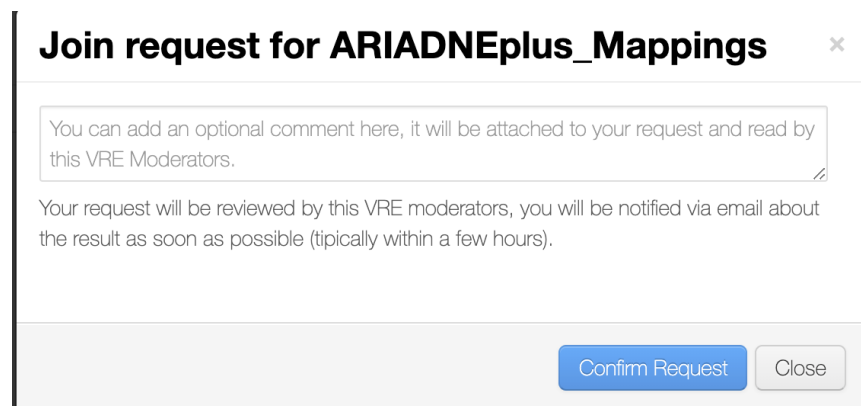
https://ariadne.d4science.org/group/ariadneplus_mappings/mapping-tool

If you are a new user you first need to register to the ARIADNEplus Mappings VRE on d4science:

- Go here <https://ariadne.d4science.org/explore>. You will see:



- Click on “Request Access” for the ARIADNEplus_Mappings
- You’ll be asked to login. If it is your first time using D4Science, you’ll be prompted with the terms of use that you have to read and accept (Click on the “I agree” button on the bottom left).
- Once logged in you’ll have to confirm the request by clicking on the “Confirm Request” button on the window that will be prompted.



- You will receive a notification email when the moderators accept your request. As soon as you receive the email you can access the environment and use the 3M mapping tool at https://ariadne.d4science.org/group/ariadneplus_mappings/mapping-tool

6.2.2. Control toolbar and Mapping list

The screenshot shows the Mapping Memory Manager interface. At the top, there are three buttons: 'New', 'View', and 'Edit'. Below these, a control toolbar contains icons for 'New', 'View', 'Edit', and 'More'. A red box highlights the 'More' icon. Below the toolbar is a 'Mappings' section with a search bar and a table of mappings. A red box highlights the table, which contains the following data:

Title	General Description	Creator	Card Status	Last Modified	M
ARIADNEplus ADS example mapping with AO-CAT 1.0.2 rdfs (1st try with a test xml)		math	Unpublished	2019-09-19	Mapping561
ARIADNEplus ADS example mapping with AO-CAT 1.0.2 rdfs (1nd try, suggested xml)		math	Unpublished	2019-09-19	Mapping571
ARIADNEplus ADS Sites and Monuments mapping (AO-CAT v1.0.2 and simplified subjects/temporal)		math	Unpublished	2019-09-20	Mapping574
ARIADNEplus ADS Sites and					

- Control toolbar for working with mappings.
- To view or edit a map, first select it in the Mapping list, then click the appropriate icon.



14

6.2.3. Control toolbar – More/Copy XML/Rights

3. Control toolbar - More

The screenshot shows the Mapping Memory Manager interface with the 'More' dropdown menu open. The menu options are:

- Publish
 - Request for publishing
- Import/Export
 - Export Template
 - Export to XML
 - Export Schema
 - Import from XML
- Versions
 - Create Version
 - View Versions
- Delete
- Unlock File
- Copy XML
- Compare
- Rights

- **More** dropdown menu gives many options for working on a mapping.
- **Copy XML** allows you to take a full copy of the selected mapping (useful when not sure of changes).



Rights allows you to share edit rights to other users.

15

6.2.4. Search and Filter

The screenshot shows the 3M Mapping Memory Manager interface. At the top, there are labels for 'Filter' and 'Search'. The 'Search' label points to a search bar in the top right corner. The 'Filter' label points to a filter dropdown menu in the top left corner. Below the search bar, there is a 'Search' button. Below the filter dropdown, there is a 'Filter Table' dropdown menu. The main content area displays a table of mappings with columns for Title, General Description, Creator, Card Status, Last Modified, and Id. The table is filtered to show 'ARIADNEplus' mappings. The sidebar on the left contains a 'Main Menu' with 'Mappings' and 'Help' sections. The 'Help' section includes 'Quick Start Guide', 'Manual', and 'X3ML Generators Manual'.

- There is a search feature allowing you to find mappings.
- There is a filter feature to narrow down within a viewed search set.



16

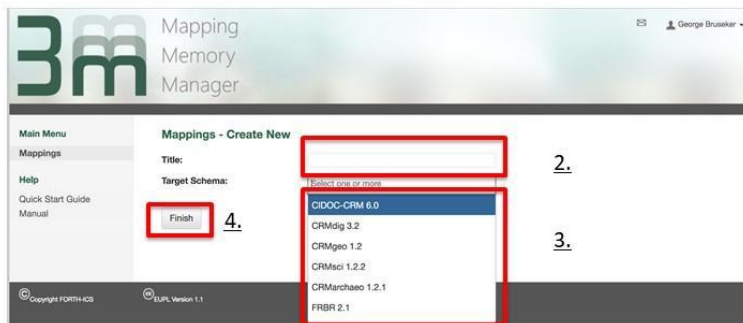
6.2.5. Manual

The screenshot shows the 3M Mapping Memory Manager interface. At the top, there is a label for 'Manual' pointing to the 'Manual' link in the sidebar. The sidebar on the left contains a 'Main Menu' with 'Mappings' and 'Help' sections. The 'Help' section includes 'Quick Start Guide', 'Manual', and 'X3ML Generators Manual'. The main content area displays a table of mappings with columns for Title, General Description, Creator, Card Status, Last Modified, and Id. The table is filtered to show 'ARIADNEplus' mappings. The search bar and filter dropdown are also visible.

- A Quick Start Guide and the complete documentation for 3M including general mapping instructions and the generators manual are always accessible directly from the sidebar.

6.3. Setting up a New Mapping

6.3.1. Give Title and choose Target Schemas



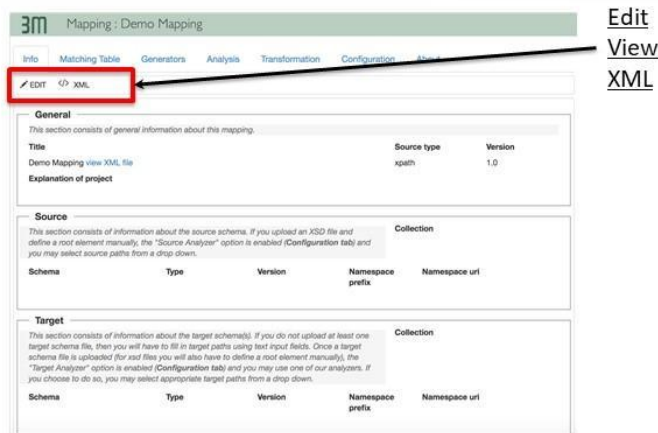
Creating a new mapping involves:

1. Click the '+' button (Create New) in 3M to open a create new map dialogue
2. Assign a title so it can be found again
3. Choose target schema(s) you wish to use
4. Finish the creation of the new map
5. Once map saved, find it again from list and open using 'edit function' (see above)



20

6.3.2. Edit/view Info Tab



- To edit the mapping, you will need to enter edit mode. This can be done by choosing the 'Edit' button.
- To leave edit mode, you can choose the opposite 'View' button. Experts can also edit the XML directly by clicking the 'XML' button.



21

6.3.3. Add Generic Mapping Metadata

Mapping Metadata provides crucial provenance information to the mapping so that it can be re-used in the future.

Who did this mapping? How can they be contacted?

Mapping

This section consists of information about who creates and supports this mapping.

Created by (Organization)

Contact person(s)

In collaboration with

6.3.4. Adding Source Schema

Upload Source Schema

Source
This section consists of information about the source schema. If you upload an XSD file and define a root element manually, the "Source Analyzer" option is enabled (Configuration tab) and you may select source paths from a drop down.

Schema

Type

Version

Collection

Adding source schema allows 3M to help guide mapping by only allowing valid paths to be selected from source.

- Upload a source schema using 'upload button'
- Source can be in XML or XSD.
- Additional metadata can be added for
 - the type of schema
 - the version of the schema
 - what collection schema used for

6.3.5. Adding Sample Data

Sample data

This section consists of information about example data (source and target) and generator policy. Once a source record XML file is uploaded, the "Transformation" tab is enabled (Transformation tab). In order to test how your source record XML file transforms to RDF/XML, N-triples or Turtle, you will probably also have to upload a generator policy XML file.

If you have not uploaded an XSD source schema yet, the "Source Analyzer" option will also be enabled once a source record XML file is uploaded (Configuration tab) and you may select source paths from a drop-down.

Provided by

Contact person(s)

Source record

Upload html

Upload xml

Generator policy

Upload xml

Target record

Upload rdf

- Upload Source Data by choosing 'Upload xml' button
- Add provenance information of who gave data and how they can be contacted.

[Upload Sample Source Data](#)

Adding source sample data allows testing of the mapping with real data using the 3M transformation tool.

Sample source record

```
1 <?xml version='1.0' encoding='UTF-8'?>
2 <Collections xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3   xmlns:dc="http://purl.org/dc/elements/1.1/"
4   xmlns:dcterms="http://purl.org/dc/terms/"
5   xmlns:ads="https://archaeologydataservice.ac.uk/">
6
7   <collection>
8     <meta>
9       <dc:title>Historic England National Record of the Historic Environment (NRHE) </dc:title>
10      <dc:identifier>1000398</dc:identifier>
11      <dc:source>Historic England</dc:source>
12      <dc:description>The Historic England National Record of the Historic Environment (NRHE) consists of core elements of the record of England&apos;s
13      archaeological and architectural sites held by the NRHE. It encompasses the historic environment in its widest sense and includes archaeological,
14      architectural and historical sites from earliest times to the present day, covering England and its territorial waters (the 12 mile limit). Each
15      record is linked to its corresponding entry in Historic England&apos;s Pastscape web site. </dc:description>
16      <dc:publisher>Archaeology Data Service</dc:publisher>
17      <dc:creator>Historic England</dc:creator>
18      <dc:date>15 Sept 2004</dc:date>
19      <dc:format>Dataset Collections</dc:format>
20      <dc:language>eng</dc:language>
21      <dc:rights>http://archaeologydataservice.ac.uk/advice/termsOfUseAndAccess</dc:rights>
22    </meta>
23    <records>
24      <record>
25        <dc:title>Part of a cropmark with...</dc:title>
26        <dc:creator>Historic England</dc:creator>
27        <dc:subjectPeriod>
28          <dc:subject>CROPMARK</dc:subject>
29          <dcterms:temporal>PALAEOLITHIC</dcterms:temporal>
30        </dc:subjectPeriod>
31        <dc:subjectPeriod>
32          <dc:subject>HOUSE</dc:subject>
33        </dc:subjectPeriod>
34        <dc:description>Part of a cropmark with internal markings seen on air photograph. Field investigation in 1957 found the marks to be due to
35        natural undulations and vegetational changes, and no traces of antiquity were observed.</dc:description>
36        <dc:identifier>Depositor ID: NT 70 NE 1</dc:identifier>
37        <dc:identifier>NMR_NATINV-3</dc:identifier>
38        <dc:source>https://archaeologydataservice.ac.uk/archsearch/record?titleId=967971</dc:source>
39        <dc:language xsi:type="dcterms:ISO639-2">eng</dc:language>
40        <dcterms:spatial>ALWINTON, ALNWICK, NORTHUMBERLAND, ENGLAND</dcterms:spatial>
41        <dcterms:spatial xsi:type="dcterms:POINT">-2.758696, 55.945225</dcterms:spatial>
42        <dc:rights>http://archaeologydataservice.ac.uk/advice/termsOfUseAndAccess</dc:rights>
43      </record>
44      <record>
45        <dc:title>Square mark seen on air...</dc:title>
46        <dc:creator>Historic England</dc:creator>
47        <dc:subjectPeriod>
48          <dc:subject>CROPMARK</dc:subject>
49          <dcterms:temporal>UNCERTAIN</dcterms:temporal>
50        </dc:subjectPeriod>
51      </record>
52    </records>
53  </collection>
54 </Collections>
```

6.3.6. Adding Target Schemas

Upload Target Schema

Target

This section consists of information about the target schema(s). If you do not upload at least one target schema file, then you will have to fill in target paths using text input fields. Once a target schema file is uploaded (for xsd files you will also have to define a root element manually), the "Target Analyzer" option is enabled (Configuration tab) and you may use one of our analyzers. If you choose to do so, you may select appropriate target paths from a drop down.

Schema **Type** **Version** **Namespace prefix** **Namespace uri**

- Click '+Target' to add new target
- Upload target schema using 'upload' button
- For each target schema specify
 - it's namespace prefix e.g.: 'aocat'
 - It's full namespace URI
- Additional metadata can be added:
 - name of the schema
 - type of schema
 - version of the schema
 - collection used for

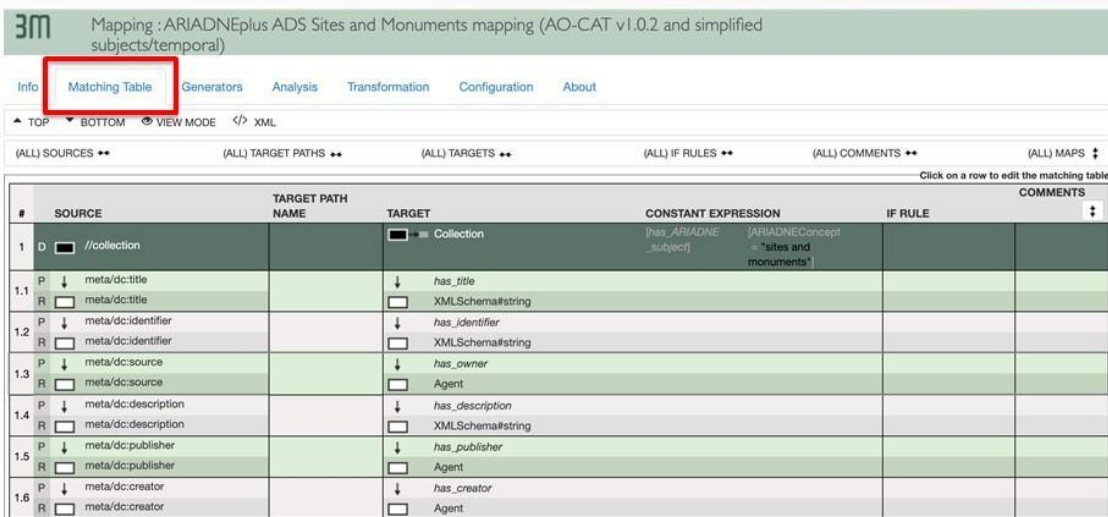
Adding target schema allows 3M to help guide mapping by only allowing valid paths to be selected from target.

Target			Collection	
<small>This section consists of information about the target schema(s). If you do not upload at least one target schema file, then you will have to fill in target paths using text input fields. Once a target schema file is uploaded (for xsd files you will also have to define a root element manually), the "Target Analyzer" option is enabled (Configuration tab) and you may use one of our analyzers. If you choose to do so, you may select appropriate target paths from a drop down.</small>				
Schema	Type	Version	Namespace prefix	Namespace uri
CIDOC-CRM view	rdfs	6.2.1	crm	http://www.cidoc-crm.org/cidoc-crm/
AO-CAT view	rdfs	1.0.2	aocat	https://www.ariadne-infrastructure.eu/resource/ao/cat/1.0/
			owl	http://www.w3.org/2002/07/owl#
			xml	http://www.w3.org/XML/1998/namespace
			skos	http://www.w3.org/2004/02/skos/core#
			crmpe	http://parthenos.d4science.org/CRMext/CRMpe/
			crm	http://www.cidoc-crm.org/cidoc-crm/
Namespaces				
<small>This section consists of information about namespaces not declared in source or target schemas block.</small>				
Namespace prefix	Namespace uri			
ariadneplus	https://ariadne-infrastructure.eu/aocat/			

In this example CIDOC-CRM and AO-CAT are loaded.

6.4. Matching Table Operation

6.4.1. Accessing the Matching Table



Mapping : ARIADNEplus ADS Sites and Monuments mapping (AO-CAT v1.0.2 and simplified subjects/temporal)

Info **Matching Table** Generators Analysis Transformation Configuration About

TOP BOTTOM VIEW MODE XML

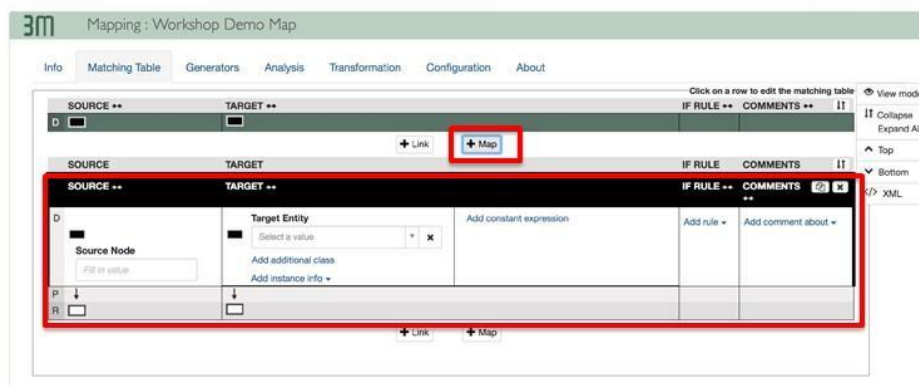
(ALL) SOURCES ** (ALL) TARGET PATHS ** (ALL) TARGETS ** (ALL) IF RULES ** (ALL) COMMENTS ** (ALL) MAPS ⌵

Click on a row to edit the matching table

#	SOURCE	TARGET PATH NAME	TARGET	CONSTANT EXPRESSION	IF RULE	COMMENTS
1	D <input type="checkbox"/> //collection		<input type="checkbox"/> Collection	{has_ARIADNE_subject} {ARIADNEConcept="sites and monuments"}		
1.1	P <input type="checkbox"/> meta/dc:title		<input type="checkbox"/> has_title			
	R <input type="checkbox"/> meta/dc:title		<input type="checkbox"/> XMLSchema#string			
1.2	P <input type="checkbox"/> meta/dc:identifier		<input type="checkbox"/> has_identifier			
	R <input type="checkbox"/> meta/dc:identifier		<input type="checkbox"/> XMLSchema#string			
1.3	P <input type="checkbox"/> meta/dc:source		<input type="checkbox"/> has_owner			
	R <input type="checkbox"/> meta/dc:source		<input type="checkbox"/> Agent			
1.4	P <input type="checkbox"/> meta/dc:description		<input type="checkbox"/> has_description			
	R <input type="checkbox"/> meta/dc:description		<input type="checkbox"/> XMLSchema#string			
1.5	P <input type="checkbox"/> meta/dc:publisher		<input type="checkbox"/> has_publisher			
	R <input type="checkbox"/> meta/dc:publisher		<input type="checkbox"/> Agent			
1.6	P <input type="checkbox"/> meta/dc:creator		<input type="checkbox"/> has_creator			
	R <input type="checkbox"/> meta/dc:creator		<input type="checkbox"/> Agent			

- From within your mapping, click on 'matching table' tab

6.4.2. Adding a New Map



Mapping : Workshop Demo Map

Info Matching Table Generators Analysis Transformation Configuration About

SOURCE ** TARGET ** IF RULE ** COMMENTS **

Click on a row to edit the matching table

+ Link **+ Map**

SOURCE **	TARGET **	IF RULE **	COMMENTS **
D <input type="checkbox"/> Source Node	Target Entity Select a value Add additional class Add instance info	Add constant expression	Add rule Add comment about
P <input type="checkbox"/>	<input type="checkbox"/>		
R <input type="checkbox"/>	<input type="checkbox"/>		

+ Link + Map

- Maps are used for mapping a domain (root node) from the source to a specific class in the target ontology
- If your data structure is complex (not flat) you can use multiple maps to map multiple domains (root nodes) in the source data structure
- Use '+ Map' button to add additional maps

Adding Map Examples

```

1 <?xml version='1.0' encoding='UTF-8'?>
2 <Collections xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3   xmlns:dc="http://purl.org/dc/elements/1.1/"
4   xmlns:dcterms="http://purl.org/dc/terms/"
5   xmlns:ads="https://archaeologydataservice.ac.uk/">
6
7   <collection>
8
9     <dc:title>Historic England National Record of
10    <dc:identifier>1000398</dc:identifier>
11    <dc:source>Historic England</dc:source>
12    <dc:description>The Historic England National
13    archaeological and architectural sites held by
14    record is linked to its corresponding entry in
15    <dc:publisher>Archaeology Data Service</dc:publisher>
16    <dc:creator>Historic England</dc:creator>
17    <dc:date>15 Sept 2004</dc:date>
18    <dc:format>Dataset Collection</dc:format>
19    <dc:language>eng</dc:language>
20    <dc:rights>http://archaeologydataservice.ac.uk
21
22    </meta>
23    <records>
24    <record>Part of a cropmark with...</dc:
25    <dc:creator>Historic England</dc:creator>
26    <dc:subjectPeriod>
27
28
29
30
31

```

- We add maps for significant nodes in the source XML that have many relations to re-express in the target ontology. These are primary foci of interest.
- In this example 'Collections' is the root node under which individual 'collection' records in this schema are documented. Each 'collection' includes
- general meta information, attributing title, identifier etc. grouped under the 'meta' node and
- a list of 'record' nodes grouped under the 'records' node.
- Our focus of interest are **collection** and **record** and each becomes the source domain of a map. The two maps will be linked eventually.
- The splitting in maps simplifies mapping and makes the mapping definition easier to read.

6.4.3. Adding a New Link

- A link is used for mapping one field/node under a domain (root node) from the source to a semantic path in the target ontology
- For each field/node in the source that should be mapped, there should be at least one link in the map
- Use '+ Link' button to add additional links

Example

```

1 <?xml version='1.0' encoding='UTF-8'?>
2 <Collections xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3   xmlns:dc="http://purl.org/dc/elements/1.1/"
4   xmlns:dcterms="http://purl.org/dc/terms/"
5   xmlns:ads="http://archaeologydataservice.ac.uk/"
6
7 * <collection>
8 *   <meta>
9     <dc:title>Historic England National Record of
10    <dc:identifier>1000398</dc:identifier>
11    <dc:source>Historic England</dc:source>
12    <dc:description>The Historic England National
13    archaeological and architectural sites held by
14    architectural and historical sites from earlie
15    record is linked to its corresponding entry in
16    <dc:publisher>Archaeology Data Service</dc:pub
17    <dc:creator>Historic England</dc:creator>
18    <dc:date>15 Sept 2004</dc:date>
19    <dc:format>Dataset Collection</dc:format>
20    <dc:language>eng</dc:language>
21    <dc:rights>http://archaeologydataservice.ac.uk
22    </meta>
23    <records>
24    <record>
25      <dc:title>Part of a cropmark with...</dc:
26      <dc:creator>Historic England</dc:creator>
27      <dc:subjectPeriod>

```

EXAMPLE:

- The domain (root node) 'collection' contains meta information about a collection in the source
 - 'collection' has several properties under the meta node: 'title', 'identifier', 'source', 'description', 'publisher', 'creator', 'date', 'format', 'language', 'rights'
- For each property we want to map, we need one link in the map

6.4.4. Copy and Deleting Maps and Links

- Longer maps/links can take time to generate. The copying feature allows one to quickly duplicate existing maps/links.
- To copy a map or link click the 'copy' icon, followed by the 'paste' icon
- To delete a map/link, select the appropriate map/link and then click the 'x'

6.4.5. How to comment

The screenshot shows the 3M editor interface for a 'Demo Mapping'. At the top, there are tabs for 'Info', 'Matching Table', 'Generators', 'Analysis', 'Transformation', 'Configuration', and 'About'. Below these are navigation controls: 'TOP', 'BOTTOM', 'VIEW MODE', and '</> XML'. A toolbar contains filters for '(ALL) SOURCES', '(ALL) TARGETS', '(ALL) IF RULES', '(ALL) COMMENTS', and '(ALL) MAPS'. The main area is a table with columns: '#', 'SOURCE', 'TARGET', 'CONSTANT EXPRESSION', 'IF RULE', and 'COMMENTS'. The first row shows a mapping from 'Source Node' to 'Target Entity'. A context menu is open over the 'COMMENTS' column, listing options: 'Add comment about', 'Rationale', 'Alternatives', 'Typical mistakes', 'Local Habits', 'Link to Cook Book', 'Example Source', 'Example Target', 'Comments Last Update Person', and 'Comments Last Update Date'.

- 3MEditor allows you to make comments on the mapping of individual fields/nodes in links. This is a useful feature to explain/remember mapping decisions
- To add comment click 'Add comment about'

6.4.6. View Controls

3MEditor also offers a series of control tools to help navigate and work with your mapping

- **Top/Bottom:** Provides scrolling over long maps
- **View Mode:** toggles map to read only
- **</> XML:** Allows direct editing of the underlying X3ML definition of the mapping

The screenshot shows the 3M editor interface for a mapping from 'ISC 19139 to CERIF'. At the top, there are tabs for 'Info', 'Matching Table', 'Generators', 'Analysis', 'Transformation', 'Configuration', and 'About'. Below these are navigation controls: 'TOP', 'BOTTOM', 'VIEW MODE', and '</> XML'. A toolbar contains filters for '(ALL) SOURCES', '(ALL) TARGETS', '(ALL) IF RULES', '(ALL) COMMENTS', and '(ALL) MAPS'. The main area is a table with columns: '#', 'SOURCE', 'TARGET', 'CONSTANT EXPRESSION', 'IF RULE', and 'COMMENTS'. The first row shows a mapping from 'Product' to 'Product'. The second row shows a mapping from 'PlainLiteral' to 'PlainLiteral'. Red boxes highlight the 'TOP', 'BOTTOM', 'VIEW MODE', and '</> XML' controls, and the '(ALL) MAPS' filter. Arrows point from these controls to the text descriptions below.

- **Collapse/Expand columns:** Hides or displays columns
- **Collapse/Expand maps:** Opens or Closes all Links in Maps

6.5. Mapping Patterns

6.5.1. X3ML Constructs

X3ML supports **1:N mappings** and uses the following special constructs:

- **intermediate nodes** used to represent the mapping of a simple source path to a complex target path.
- **constant expression nodes** used to assign constant attributes to an entity.
- **conditional statements** within the target node and target relation support checks for existence and equality of values and can be combined into Boolean expressions.
- **“Same as” variable** used to identify a specific node instance for a given input record that is generated once but is used in a number of locations in the mapping.
- **Join operator (==)** used in the source path to denote relational database joins
- **info and comment blocks** throughout the mapping specification bridge the gap between human author and machine executor.

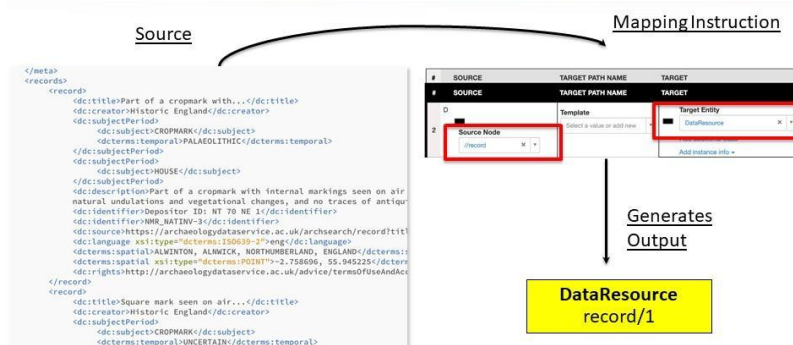
6.5.2. Mapping Source Root to Target Domain

#	SOURCE	TARGET PATH NAME	TARGET
D		Template	
2	<div style="border: 1px solid red; padding: 2px;"> Source Node <input type="text" value="//record"/> </div>	<input type="text" value="Select a value or add new"/>	<div style="border: 1px solid red; padding: 2px;"> target Entity <input type="text" value="DataResource"/> </div>

- Click top row in a map
- On Source Side, choose appropriate domain (a node) from source
- On Target Side, choose appropriate class in target ontology

This now says:

“for each instance of source domain (node) generate one instance of target class”

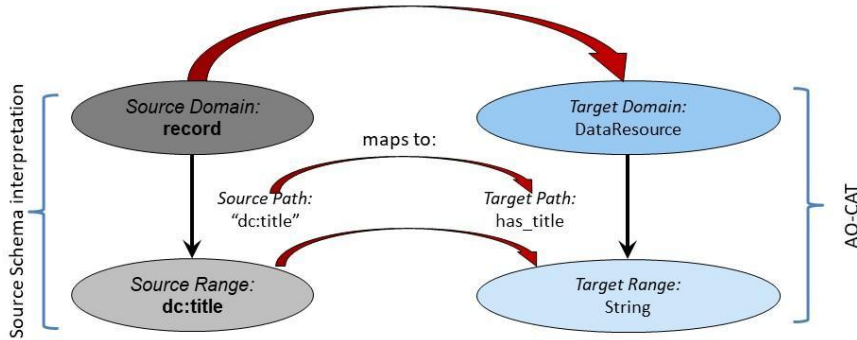


So this mapping instruction says,

“for each instance of ‘record’ in source, create an RDF instance of type ‘DataResource’ of the AO-CAT”.



Simple Field Mapping: Creating and Equivalent Proposition



6.5.3. Simple Field Mapping (One to One)

#	SOURCE	TARGET PATH NAME	TARGET
2	D <input type="checkbox"/> ../record		<input type="checkbox"/> DataResource
#	SOURCE	TARGET PATH NAME	TARGET
2	D <input type="checkbox"/> ../record		<input type="checkbox"/> DataResource
P	Source Relation dc:title x	Template Select a value or add new	Target Relation has_title x
2.1	R Source Node dc:title x		Target Entity XMLSchema#string x

- Choose field for mapping from Source
- Add one relation and one target class in target
- This creates a simple triple statement in RDF of form S-P-O

Example:

Source

```

</meta>
<record>
  <dc:title>Part of a cropmark with...</dc:title>
  <dc:spatial>https://archaeologydataservice.ac.uk/archsearch/record/ttitle
  <dc:subject>CROPMARK</dc:subject>
  <dc:temporal>PALAEOETHIC</dc:temporal>
  </dc:subjectPeriod>
  <dc:subject>HOUSE</dc:subject>
  </dc:subjectPeriod>
  <dc:description>Part of a cropmark with internal markings seen on air
  natural undulations and vegetational changes, and no traces of antiqu
  <dc:identifier>Depositor ID: NT 70 NE 1</dc:identifier>
  <dc:identifier>NMR_NATINV-3</dc:identifier>
  <dc:source>https://archaeologydataservice.ac.uk/archsearch/record/ttitle
  <dc:language>xsi:type="dcterms:ISO639-1">eng</dc:language>
  <dc:spatial>ALWINTON, ALNWICK, NORTHUMBERLAND, ENGLAND</dc:spatial>
  <dc:terms>spatial xsi:type="dcterms:PO2IN1">-2.758896, 55.945225</dc:terms>
  <dc:rights>http://archaeologydataservice.ac.uk/advice/termsOfUseandLic
  </dc:rights>
  </record>
  <record>
  <dc:title>Square mark seen on air...</dc:title>
  <dc:spatial>https://archaeologydataservice.ac.uk/archsearch/record/ttitle
  <dc:subject>CROPMARK</dc:subject>
  <dc:temporal>UNCERTAIN</dc:temporal>
  </dc:subjectPeriod>
  </record>
        
```

Mapping Instruction

#	SOURCE	TARGET PATH NAME	TARGET
2	D <input type="checkbox"/> ../record		<input type="checkbox"/> DataResource
#	SOURCE	TARGET PATH NAME	TARGET
2	D <input type="checkbox"/> ../record		<input type="checkbox"/> DataResource
P	Source Relation dc:title x	Template Select a value or add new	Target Relation has_title x
2.1	R Source Node dc:title x		Target Entity XMLSchema#string x

Generates Output

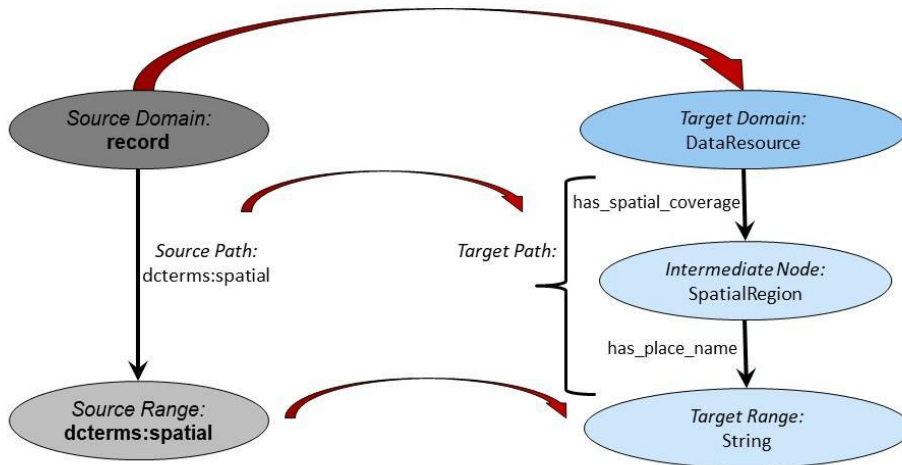
```

DataResource Record/1 has_title 'Part of a cropmark with ...'
DataResource Record/2 has_title 'Square mark seen on air ...'
        
```

So this mapping instruction says,

“for each instance of ‘dc:title’ in ‘record’, create a RDF triple stating that this instance of ‘DataResource’ has title the instance string”

6.5.4. Mapping to Paths: Introducing intermediate nodes



#	SOURCE	TARGET PATH NAME	TARGET	CONSTANT EXPRESSION	IF RULE	COMMENTS
2	record		DataResource			
P	Source Relation dcterms:spatial	Template Select a value or add new	Target Relation has_spatial_coverage	Add constant expression	NOT Existence {xsi:type}	Add comment about
			Target Entity SpatialRegion			
			Add additional class is same as (map)			
			is			
			Add instance info			
			Target Relation has_place_name			
			Add intermediate			
R	Source Node dcterms:spatial		Target Entity XMLSchemaString	Add constant expression	Add rule	Add comment about
			Add additional class			
			Add instance info			

For more complex paths, intermediate nodes are necessary.

- 'Click 'Add intermediate' for as many intermediate relations as you may need to build.

In this example, the place name of the SpatialRegion is expressed.

Example:

Source

```

</meta>
<records>
<record>
<dc:title>Part of a cropmark with ...</dc:title>
<dc:creator>Historic England</dc:creator>
<dc:subjectPeriod>
<dc:subject>CROPHARK</dc:subject>
<dc:temporal>PALAEOETHIC</dc:temporal>
</dc:subjectPeriod>
<dc:subjectPeriod>
<dc:subject>HOUSE</dc:subject>
</dc:subjectPeriod>
<dc:description>Part of a cropmark with internal markings seen on air
natural undulations and vegetational changes, and no traces of antiqu
<dc:identifier>Depositor ID: NT 70 NE 1</dc:identifier>
<dc:identifier>MNR_NATINW-3</dc:identifier>
<dc:source>https://archaeologydataservice.ac.uk/archsearch/record?tit
<dc:temporal>NT 70 NE 1</dc:temporal>
<dc:spatial>ALWINTON, ALNWICK, NORTHUMBERLAND, ENGLAND</dc:spatial>
<dc:spatial>{xsi:type="dcterms:SPATIAL"}>753696
<dc:rights>http://archaeologydataservice.ac.uk/adv/
</record>
<record>
<dc:title>Square mark seen on air...</dc:title>
<dc:creator>Historic England</dc:creator>
<dc:subjectPeriod>
<dc:subject>CROPHARK</dc:subject>
<dc:temporal>UNCERTAIN</dc:temporal>
          
```

Mapping Instruction

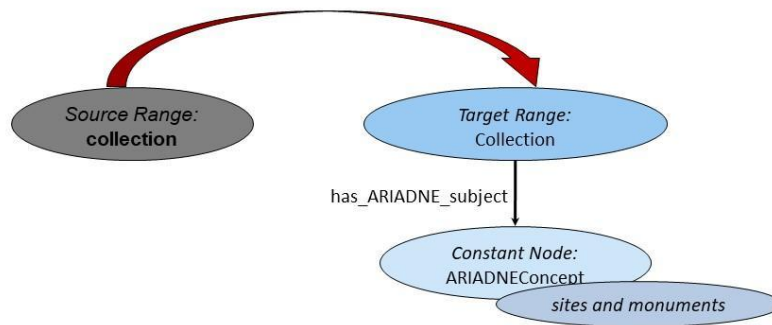
Generates Output

```

DataResource Row/1 -- has_spatial_coverage --> SpatialRegion
SpatialRegion -- has_name --> String
ALWINTON, NORTHUMBERLAND, ENGLAND
          
```

So this mapping instruction says, "for each instance of 'dcterms:spatial' in 'record', create a series of RDF triples stating that 'DataResource' has spatial coverage 'SpatialRegion' has place name "String"

6.5.5. Mapping to Paths: Introducing constant expressions



#	SOURCE	TARGET PATH NAME	TARGET	CONSTANT EXPRESSION
1	Source Node //collection	Template Select a value or add new	Target Entity Collection	CONSTANT EXPRESSION Relation has_ARIADNE_subject Entity ARIADNEConcept Constant sites and monuments

Sometimes, we want to add additional constant information to our mapping.

- Click 'Add Constant Expression' (to the right of target)
- Add appropriate links and constant info

In this example, we want to express for each node representing a collection in the target graph has ARIADNE subject 'sites and monuments'.

Example:

Source

```

1 <?xml version='1.0' encoding='UTF-8'?>
2 <Collections xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance'
3   xmlns:dc='http://purl.org/dc/elements/1.1/'
4   xmlns:dcterms='http://purl.org/dc/terms/'
5   xmlns:ads='http://archaeologydataservice.ac.uk/'?>
6
7   <collection>
8     <title>Historic England National Record of
9     <dc:identifier>1000398</dc:identifier>
10    <dc:source>Historic England</dc:source>
11    <dc:description>The Historic England National
12    archaeological and architectural sites held by
13    record is linked to its corresponding entry in
14    <dc:publisher>Archaeology Data Service</dc:pub
15    <dc:creator>Historic England</dc:creator>
16    <dc:date>15 Sept. 2004</dc:date>
17    <dc:format>Dataset</dc:format>
18    <dc:language>eng</dc:language>
19    <dc:rights>http://archaeologydataservice.ac.uk
20  </meta>
21  <records>
22    <record>
23      <dc:title>Part of a cropmark with...</dc:
24      <dc:creator>Historic England</dc:creator>
25      <dc:subject>Period
          
```

Mapping Instruction

TARGET	CONSTANT EXPRESSION
Target Entity Collection	CONSTANT EXPRESSION Relation has_ARIADNE_subject Entity ARIADNEConcept Constant sites and monuments

Generates Output

Collection
Collection/1000398

has_ARIADNE_subject

ARIADNEConcept
sites and monuments

So this mapping instruction says, "for each instance of 'collection', create a series of RDF triples stating that: 'Collection' has_ARIADNE_subject 'ARIADNEConcept' 'sites and monuments'"

6.5.6. Using variables

Add instance info ▾

Is same as (map)
 Is same as (global)

Source Relation	Target Relation
dcterms:spatial	has_spatial_coverage
<input type="checkbox"/> Add intermediate <input type="checkbox"/> Set blank source relation	<input type="checkbox"/> SpatialRegion <input checked="" type="radio"/> Is same as (map) <input type="text" value="sr1"/>
	<input type="checkbox"/> Add instance info ▾ <input type="checkbox"/> Add intermediate

When two nodes in a mapping refer to the same real world entity, we can tell the mapping engine to generate only one node for this entity

For each node we wish to merge,

- click 'Add instance info'
- Click 'is same as'
- Add variable (random name) to both nodes

2.9	P	dcterms:spatial	has_spatial_coverage	SpatialRegion [sr1]
	R	dcterms:spatial	has_place_name	XMLSchema#string
2.10	P	dcterms:spatial	has_spatial_coverage	SpatialRegionPoint [sr1]
	R	dcterms:spatial	has_longitude	XMLSchema#decimal

In this example the variable 'sr1' denotes that for each record, the instance of SpatialRegion is the same with the instance of SpatialRegionPoint. Note that SpatialRegionPoint is a subclass of SpatialRegion.



6.5.7. Joining Maps

#	SOURCE	TARGET PATH NAME	TARGET
1	D	./collection	Collection
1.1	P	./dc:title	has_title
	R	./dc:title	XMLSchema#string
1.2	P	./dc:identifier	has_identifier
	R	./dc:identifier	XMLSchema#string
1.3	P	./dc:source	has_owner
	R	./dc:source	Agent
1.4	P	./dc:description	has_description
	R	./dc:description	XMLSchema#string
1.5	P	./dc:publisher	has_publisher
	R	./dc:publisher	Agent
1.6	P	./dc:creator	has_creator
	R	./dc:creator	Agent
1.7	P	./dc:issued	was_issued
	R	./dc:issued	XMLSchema#dateTime
1.8	P	./dc:format	has_type
	R	./dc:format	ARIADNEConcept
1.9	P	./dc:language	has_language
	R	./dc:language	XMLSchema#string
1.10	P	./dc:rights	has_access_rights
	R	./dc:rights	XMLSchema#string
1.11	P	./record	has_part
	R	./record	DataResource

#	SOURCE	TARGET PATH NAME	TARGET
2	D	./record	DataResource
2.1	P	./dc:title	has_title
	R	./dc:title	XMLSchema#string
2.2	P	./dc:creator	has_creator
	R	./dc:creator	Agent
2.3	P	./dc:subject	has_native_subject
	R	./dc:subject	ARIADNEConcept
2.4	P	./dcterms:temporal	has_temporal_coverage
	R	./dcterms:temporal	TemporalRegion
2.5	P	./dc:description	has_description
	R	./dc:description	XMLSchema#string
2.6	P	./dc:identifier	has_identifier
	R	./dc:identifier	XMLSchema#string
2.7	P	./dc:source	has_landing_page
	R	./dc:source	Thing
2.8	P	./dc:language	has_language
	R	./dc:language	XMLSchema#string
2.9	P	./dcterms:spatial	has_spatial_coverage
	R	./dcterms:spatial	SpatialRegion [sr1]
	R	./dcterms:spatial	has_place_name
	R	./dcterms:spatial	XMLSchema#string
	R	./dcterms:spatial	has_spatial_coverage

Two maps can be joined, if the first maps a node which is the domain of the second map. This is often done for convenience purposes, to make mapping easier and simple to read.

E.g.: in the ADS example, 'collection' has a nested list of 'record' nodes. Each 'record' node has many fields describing the record. In map 1, collection references the 'record'. In map 2, we map the fields for 'record' node.



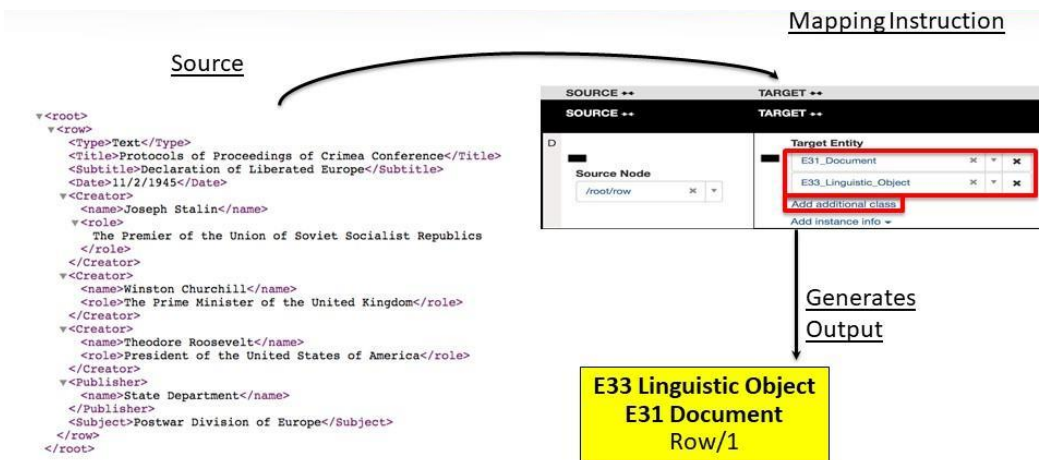
6.5.8. Multiple instantiation

SOURCE ↔	TARGET ↔								
SOURCE ↔ D Source Node <input type="text" value="/root/row"/>	TARGET ↔ Target Entity <table border="1"> <tr> <td>E31_Document</td> <td>x</td> <td>v</td> <td>x</td> </tr> <tr> <td>E33_Linguistic_Object</td> <td>x</td> <td>v</td> <td>x</td> </tr> </table> <input type="button" value="Add additional class"/> <input type="button" value="Add instance info"/>	E31_Document	x	v	x	E33_Linguistic_Object	x	v	x
E31_Document	x	v	x						
E33_Linguistic_Object	x	v	x						

For any node, if it is an instance of two or more classes, we can express this in 3M by using the 'add additional class' feature.

In this example, we say each 'row' from source both is an E31 Document and an E33 Linguistic Object. This allows the use of relations from both classes.

Example:



So this mapping instruction says,
 “for each instance of ‘row’ in source, create an RDF instance of
 both type ‘E31 Document’ and ‘E33 Linguistic Object’”



6.5.9. Mapping under conditions - If rules

#	SOURCE	TARGET PATH NAME	TARGET	CONSTANT EXPRESSION	IF RULE	COMMENTS
2	Source Relation Add intermediate Set blank source relation	Template Select a value or add new	Target Relation has_spatial_coverage Target Entity SpatialRegion Add additional class is same as (map) xsi Add instance info Target Relation has_place_name Add intermediate	Add constant expression	NOT Existence xsi:type Add rule	Add comment about
2.9	Source Node xsi:type		Target Entity XMLSchemaSetting Add additional class Add instance info	Add constant expression	Add rule	Add comment

For any node, we can express one or more conditions and the mapping will happen only if the conditions are met.
In this example, we check if the attribute `xsi:type` does NOT exist in the source file.



6.6. Instance and label Generators

6.6.1. Why Instance and Label Generators, what do they do?

The mapping table allows you to make a translation between a source schema and an RDFS encoded schema like CIDOC CRM. Each node specified in the target will become a separate data entity in the semantic graph that is created through the X3ML transformation engine. This separate data entity will need a unique identifier by which it can be found in the system (like a unique key in a relational database). The instance generators allow specifying patterns for building unique identifiers for instances called 'URIs'.

Because a URI is often unreadable, it is highly recommended that for each node a label (usually the actual data value from your source schema) be added to each node as well. This is also done through the instance generator.

6.6.2. Defining instance generation functions/patterns (offline)

The instance generator is a small piece of simple XML code. This is generated with a simple XML file outside X3ML defining convenient patterns. It can then be uploaded to 3MEditor and used to assign the instance generators.

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <generator_policy>
3   <!-- The prefix is a namespace that must be declared in the X3ML definition.
4   in the x3ml define <namespace prefix="ariadneplus" uri="http://www.ariadneplus.org" />
5   <generator name="CollectionURI" prefix="ariadneplus">
6     <pattern>Collection/{term}</pattern>
7   </generator>
8   <generator name="DataResourceURI" prefix="ariadneplus">
9     <pattern>Resource/{term}</pattern>
10  </generator>
11  <generator name="IdentifierURI" prefix="ariadneplus">
12    <pattern>Identifier/{term}</pattern>
13  </generator>
14  <generator name="EventURI" prefix="ariadneplus">
15    <pattern>Event/{term}</pattern>
16  </generator>
17  <generator name="Time-SpanURI" prefix="ariadneplus">
18    <pattern>Time-Span/{term}</pattern>
19  </generator>
20  <generator name="PlaceURI" prefix="ariadneplus">
21    <pattern>Place/{term}</pattern>
22  </generator>
23  <generator name="AgentURI" prefix="ariadneplus">
24    <pattern>Agent/{term}</pattern>
25  </generator>
26  <generator name="ConceptURI" prefix="ariadneplus">
27    <pattern>Concept/{term}</pattern>
28  </generator>
29  <generator name="ConceptURI-2" prefix="ariadneplus">
30    <pattern>Concept/{scheme}/{term}</pattern>
31  </generator>
32  <generator name="OneLevelCustomURI" prefix="ariadneplus">
33    <pattern>{level1}/{term}</pattern>
34  </generator>
35  <generator name="TwoLevelCustomURI" prefix="ariadneplus">
36    <pattern>{level1}/{level2}/{term}</pattern>
37  </generator>
38  <generator name="SimpleLabel">
39    <pattern>{label}</pattern>
40  </generator>
41  <generator name="CompositeLabel">
42    <pattern>{label} {text}</pattern>
43  </generator>

```

Prefix will be used to define base path of URI



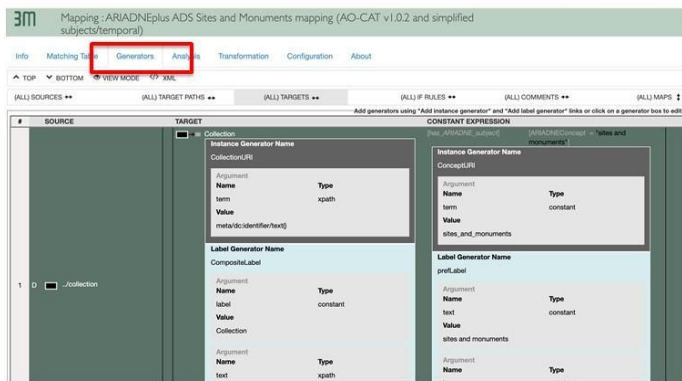
6.6.3. Adding generator file (info tab)

Namespaces		
Namespace prefix	Namespace uri	
ariadneplus	https://ariadne-infrastructure.eu/locat/	
+ Namespace		
Mapping		
This section consists of information about who creates and supports this mapping.		
Created by (Organization)	Contact person(s)	In collaboration with
<input type="text"/>	<input type="text"/>	<input type="text"/>
Sample data		
This section consists of information about example data (source and target), generator policy and thesaurus. Once a source record XML file is uploaded, the "Transformation" tab is enabled (Transformation tab). In order to test how your source record XML file transforms to RDF/XML, N-triples or Turtle, you will probably also have to upload a generator policy XML file. You may also want to upload a thesaurus file. If you have not uploaded an XSD source schema yet, the "Source Analyzer" option will also be enabled once a source record XML file is uploaded (Configuration tab) and you may select source paths from a drop down.		
Provided by	Source record	Target record
<input type="text"/>	ariadne_testSample.xml	<input type="text"/>
Contact person(s)	Generator policy	Thesaurus
<input type="text"/>	ARIADNEplusGeneratorPole	<input type="text"/>
	<input type="button" value="view generator xml"/>	<input type="button" value="view target"/>
		<input type="button" value="Upload file"/>

1. Add a new namespace, and specify the same namespace prefix as in your generator file. Use 'namespace uri' field to specify your base namespace.
2. Upload the generator file.



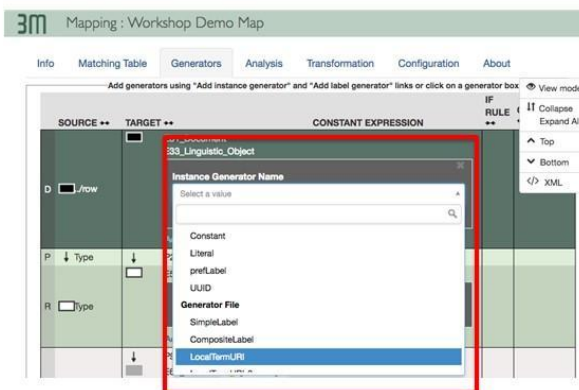
6.6.4. Opening up the generator editor



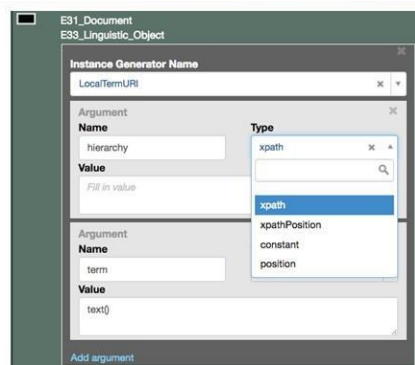
- Opening generator editor is easy, just click to the generators tab.
- The tool adds by default the UUID generator to all entities for fast testing.



6.6.5. Specifying Instance Generators



1. Click Add Instance Generator
2. Select Instance Generator



1. Each argument in a generator specifies a piece of the URI string
2. Each argument in the URI can be of different types, chiefly 'Constant' (a value you specify) or 'xpath' (a value from the source)
3. The unique portion of the URI should usually be drawn from the source data. It is accessed from a node using the command 'text()'

6.6.6. Testing Transforms

Mapping :ARIADNEplus ADS Sites and Monuments mapping (AO-CAT v1.0.2 and simplified subjects/temporal)

Info Matching Table Generators Analysis Transformation Configuration About

Metadata Transformation

If you have a Generator Policy xml file available, please upload it (Info tab). Otherwise you are provided with a generic one. If a target record file has already been uploaded (Info tab), you may view it in "Target Record RDF File".

You may modify both source and generator files on the fly, but changes will be lost next time you click on the Transformation tab. (You may read instructions about generators by clicking [Notes about Generators](#)). Check the ADS site and output format and their class. "Run Transformation". View metadata transformed for client application managed mode "Target Record RDF File" sidebar. You may save transformed metadata as your target record file by clicking "Save as Target Record" button. However, this action will overwrite any previous metadata, so use it with caution!

ATTENTION: This version of the editor produces client mapping files compatible with client engine version 1.0.0 or newer. The mapping files are not backward compatible with older versions of the client engine. Download client engine at [SIP-ICE](#).

Source Record XML File

```
<?xml version="1.0" encoding="UTF-8"?><Collection xmlns:xsi="https://www.w3.org/2001/XMLSchema-instance" xmlns="http://purl.org/dc/terms/1.1" xsi:schemaLocation="http://purl.org/dc/terms/ http://www.w3.org/ns/xsi/2001-09-05/collect"><collection><site><site><site><site></site></site></collection></Collection>
```

Generator Policy XML File

```
<?xml version="1.0" encoding="UTF-8"?><generator-policy><!-- The prefix is a namespace that must be declared in the XML definition in the client definition --><generator name="CollectionURI" prefix="xsi:collect"><generator name="CollectionURI" prefix="xsi:collect"><generator name="CollectionURI" prefix="xsi:collect"><generator name="CollectionURI" prefix="xsi:collect"></generator></generator-policy>
```

Run Transformation

Save as Target Record

Subject

RDF in error

Visualize selected subject

To test transform all mapped entities must have an instance generator specified.

- Click on **Transform** tab
- Click **Run Transformation**
- Analyze resulting RDF and/or error messages in Transformation console

6.6.7. Visualize transformed records

- To visualize a transformed record, you must first **Save as target record**.
- Select the URI of a record, it automatically copies in the **Subject** text box
- Click **Visualize selected subject**

6.6.8. RDF visualizer

RDF Visualizer

Model Overview

Enter subject

Choose a Template: Template 1: Mark same instances

Record Robert Muschamp (living in 1212,...) [DataResource]

- has_temporal_coverage: 1212 [TemporalRegion]
- has_title: Robert Muschamp (living in 1212,...)@en
- has_native_subject: CHANEL@en [ARIADNEConcept]
- has_description: Robert Muschamp (living in 1212, died 1250) gave tithes to the Priory of Kirkham and church at Kirknewton in return for permission to have a chantry in the chapel at Hethpool. There is no known location for a chapel there.@en
- has_creator: Historic England [Agent]
- has_identifier: Depositor ID: NT 82 NE 51
- has_language: NMR_NATIVE-806

7. Data Mapping

The integrated metadata will be searchable via the ARIADNEplus portal according to three main facets: Where (place) / When (time) / What (subject). We need to support effective cross-search of metadata originating from many different countries; this requires some shared common understanding of the meaning of the metadata. The where and when facets may be communicated using common data types and comparable values (e.g. spatial coordinates relative to a known location on the earth, date ranges relative to a known epoch). The what (subject) aspect can be more difficult to define in a commonly understood and comparable format.

Clean and consistent metadata is a prerequisite for effective integration and reuse, as we can neither predict nor dictate the way in which it may be reused in the future. Inconsistent or erroneous metadata causes problems in accurately identifying co-references or mappings. Embedded errors may resurface and be magnified where the metadata is subsequently reused in a different context, causing difficulties in performing logical reasoning – and leading to potentially inaccurate assertions or conclusions in future studies.

7.1. Getty AAT

The Getty Art & Architecture Thesaurus (AAT) provides concepts and terms to describe cultural heritage concepts. AAT is [made available as Linked Open Data \(LOD\)](#), so each concept in the thesaurus has a unique identifier in the form of a URI (e.g. <http://vocab.getty.edu/aat/300211545> is the identifier of the concept "*penannular brooches*"). ARIADNEplus is using AAT as a central hub for scalable interconnection of local subject vocabularies. Data providers are required to provide a set of mappings from their own local subject vocabulary (all terms used to describe subjects in their own metadata) to the AAT. The aim of the subject mapping exercise is to improve recall and precision when later browsing and searching the ARIADNEplus integrated data by subject (time and space are being handled separately).

penannular brooches

Source: <http://vocab.getty.edu/aat/300211545>

[Website](#) | [Hierarchy](#) | Download in: [JSON](#) | [JSONLD](#) | [RDF](#) | [N3/Turtle](#) | [N-Triples](#)

Inference

Statements in which the resource exists as a subject.

Predicate	Object
rdf:type	gvp:Concept
rdfs:seeAlso	http://www.getty.edu/vow/AATFullDisplay?find=&logic=AND&note=&subjectid=300211545
dcterms:created	1992-03-05T00:00:00
dcterms:modified	2001-07-26T22:03:06, 2010-04-23T05:52:09, 2010-04-23T05:52:10, 2012-03-14T15:05:29, 2012-03-16T14:11:48, 2014-12-05T08:40:23, 2014-12-05T08:43:10
skos:changeNote	aat_rev:5000052642, aat_rev:5001133770, aat_rev:5001133771, aat_rev:5001133772, aat_rev:5002479845, aat_rev:5002479851, aat_rev:5003174190, aat_rev:5003174193, aat_rev:5003178253, aat_rev:5003763847, aat_rev:5003763848, aat_rev:5003765784
gvp:parentString	brooches, pins (jewelry), jewelry worn on costume, <jewelry by location>, jewelry, worn costume accessories, costume accessories, costume (mode of fashion), Costume (hierarchy name), Furnishings and Equipment (hierarchy name), Objects Facet
gvp:parentStringAbbrev	brooches, pins (jewelry), ... Objects Facet
gvp:displayOrder	2
xl:prefLabel	aat_term:1000211545-en, aat_term:1000538040-es, aat_term:1000611650-nl
xl:altLabel	aat_term:1000226276-en, aat_term:1000294365-en, aat_term:1000447710-es, aat_term:1000611651-nl
gvp:broaderGeneric	aat:300045995
gvp:broaderPreferred	aat:300045995
gvp:prefLabelGVP	aat_term:1000211545-en
skos:inScheme	aat
skos:scopeNote	aat_scopeNote:124494, aat_scopeNote:36428, aat_scopeNote:75533
dcterms:contributor	aat_contrib:10000000, aat_contrib:10000088, aat_contrib:10000131
dcterms:source	aat_source:2000005951-subject-300211545, aat_source:2000051089-subject-300211545
dc:identifier	300211545
dcterms:license	http://opendatacommons.org/licenses/by/1.0/
cc:license	http://opendatacommons.org/licenses/by/1.0/
void:inDataset	http://vocab.getty.edu/dataset/aat
prov:wasGeneratedBy	aat_rev:5000052642

Mappings are inherently reusable as they describe semantic relationships between persistent concepts relating to the domain, but are not restricted in scope to any particular dataset. Therefore if your organization participated in the first ARIADNE project it is likely that a set of subject mappings were produced already, in this case it may be just a matter of locating, reviewing and revising them as appropriate rather than starting again from scratch.

The mappings can be produced using the [Vocabulary Matching Tool \(VMT\)](#). The VMT is a browser-based application with no installation or configuration requirements. It has a multilingual user interface – at the time of writing the current UI languages supported are German, English, Spanish, French, Italian and Dutch. Users can search and browse the AAT, viewing the hierarchical context and scope of concepts to make a better-informed match. [Further instructions and general guidance](#) are available online.

Vocabulary Matching Tool

English

Source Concept	Match Type	Target Concept	Suggest	Delete Row
Identifier		Filter column...		
Label				
Filter column...				
http://pur.org/heritagedata/schemes... Abbey Church	Close Match	abbey churches	🔍	🗑️
http://pur.org/heritagedata/schemes... ABBEY	Exact Match	abbey (monasteries)	🔍	🗑️
http://pur.org/heritagedata/schemes... AGRICULTURAL BUILDING	Exact Match	agricultural buildings	🔍	🗑️
http://pur.org/heritagedata/schemes... AGRICULTURAL DWELLING	Broad Match	agricultural buildings	🔍	🗑️
http://pur.org/heritagedata/schemes... AGRICULTURAL HALL	Broad Match	agricultural buildings	🔍	🗑️
http://pur.org/heritagedata/schemes... FARM BUILDING	Close Match	agricultural buildings	🔍	🗑️
http://pur.org/heritagedata/schemes... FIELD SYSTEM	Broad Match	agricultural land	🔍	🗑️
http://pur.org/heritagedata/schemes... FIELD SYSTEM	Broad Match	agricultural land	🔍	🗑️
http://pur.org/heritagedata/schemes... LAND USE SITE	Broad Match	agricultural land	🔍	🗑️
http://pur.org/heritagedata/schemes... LYNCHET	Broad Match	agricultural land	🔍	🗑️
http://pur.org/heritagedata/schemes... CURVILINEAR ENCLOSURE	Broad Match	agricultural settlements	🔍	🗑️
http://pur.org/heritagedata/schemes... DITCHED ENCLOSURE	Broad Match	agricultural settlements	🔍	🗑️
http://pur.org/heritagedata/schemes... DOUBLE DITCHED ENCLOSURE	Broad Match	agricultural settlements	🔍	🗑️
http://pur.org/heritagedata/schemes... ENCLOSED SETTLEMENT	Broad Match	agricultural settlements	🔍	🗑️
http://pur.org/heritagedata/schemes... ENCLOSURE	Broad Match	agricultural settlements	🔍	🗑️
http://pur.org/heritagedata/schemes... AGRICULTURE AND SUBSISTENCE	Broad Match	agriculture	🔍	🗑️
http://pur.org/heritagedata/schemes... AIR RAID SHELTER	Exact Match	air raid shelters	🔍	🗑️
http://pur.org/heritagedata/schemes... AIRCRAFT	Close Match	aircraft	🔍	🗑️

390 rows



Created by University of South Wales

ARIADNEplus is a Horizon 2020 project funded by the European Commission (Grant Agreement No 823924)

This application retrieves some information originating from Getty Art & Architecture Thesaurus (AAT)® which is made available under the ODC Attribution License. See <http://vocab.getty.edu/> for further details.

Usually you will just make one match from any given local concept to an AAT concept - having mapped to an AAT target concept there is no need for additional mappings to narrower AAT concepts. An exception to this may be the situation where the local source concept is a composite or pre-coordinated concept related to multiple genuinely different AAT concepts e.g. “bronze spearheads”. In this case, two mappings could be made – firstly to [aat:300010957](http://aat.getty.edu/300010957) “bronze (metal)” and secondly to [aat:300037118](http://aat.getty.edu/300037118) “spearheads”.

The type of match between the source concept and the target concept will be one of the possible SKOS mapping properties as defined by the [SKOS reference document](#), listed here in order of preference:

- Exact Match - this match type indicates that there is "a high degree of confidence that the concepts can be used interchangeably across a wide range of information retrieval applications".
- Close Match - can be used to link concepts "that are sufficiently similar that they can be used interchangeably in some information retrieval applications". It is a more approximate type of relationship between the source and target concepts.
- Broad Match - expresses a hierarchical generic relationship between concepts. If a source concept is more specific in scope than any AAT available concept then you can make a Broad Match link to the AAT concept.

This is useful for cases when a source vocabulary contains concepts that are more detailed. In this case the "all/some" rule should hold true e.g. "coffee cups" Broad Match "cups" - ALL coffee cups are cups; SOME cups are coffee cups.

- Narrow Match - expresses a hierarchical generic relationship between concepts. It is not expected that you would make much use of Narrow Match relationships for ARIADNEplus vocabularies. The "some/all" rule should hold true e.g. "cups" Narrow Match "coffee cups" - SOME cups are coffee cups; ALL coffee cups are cups.
- Related Match - expresses an associative relationship between concepts. The exact nature of the relationship is not explicitly spelled out, only that there is some see also connection between them. E.g. bullets Related Match guns. Bullets are associated with guns, though you can see that this relationship is different to the other relationship types listed. Preferably you would aim to create a more direct concept mapping wherever the target AAT vocabulary supports it (in this case to [aat:300201694](#) "bullets").

Once created the sets of mappings are to be exported from the VMT and deposited on D4Science under "VRE Folders > ARIADNEplus_Mappings > Matched Vocabularies". There is a separate folder per data provider – the mappings will be subsequently picked up from here and used in the data enrichment process.

7.2. Period0

Perio.do is a multilingual gazetteer of period definitions for describing historical and archaeological named periods. It comprises a series of 'collections' (lists of named periods) where each individual period definition is associated with a particular geographical area and a start date and end date. Periods and collections have 'permalink' identifiers (URIs) so they can be clearly and unambiguously referenced - e.g. the identifier <http://n2t.net/ark:/99152/p0rrjd9gix9> defines the period "Moyen Âge" (Middle Ages) within the collection <http://n2t.net/ark:/99152/p0rrjd9> "INRAP: Chronologie Generale. 2007". All named periods in ARIADNEplus will be defined with reference to periods within Perio.do. This may necessitate creating and submitting a new collection via the Perio.do client application, however If your organization participated in the first ARIADNE project it is likely that a Perio.do collection for the periods defined in your data already exists. In this case it may be just a matter of reviewing and revising them as appropriate rather than starting from scratch. Data providers will inform the data ingestion team of the URI of the Perio.do collection to be used in interpreting named periods in the context of their dataset.

NB It is assumed that dates provided, either via Perio.do or as dates already expressed in absolute years in partner's datasets, will be expressed in years Common Era (CE)

I.e. 2020 CE or AD 2020, or 400 BCE or 400 BC

If partners hold data in years BP (using either 1950 or 2000 as present) it is preferred that this is first transformed to years CE.

Canonical / INRAP: INRAP: Chronologie Generale. 2007.

INRAP. INRAP: Chronologie Generale. 2007.

Permalink: <http://n2t.net/ark:/99152/p0rjd9>

Title INRAP: Chronologie Generale
Citation INRAP: Institut National de Recherches Archeologiques Preventives
URL <http://www.inrap.fr/archeologie-preventive/Ressources/Dossiers-multimedias/Chronologie/Chronologie-des-periodes-de-l-histoire-et-de-l-archeologie/p-12505-Chronologie-Interactive800-000-ans-eclaires-par-l-archeologie-version-texte-htm>
Year published 2007
Creators INRAP

Editorial note

Source uses "avant notre ère" interchangeably with ISO8601 dates in the "-" form. Definitions attributed to individual authors as available.

Period list

Viewing 1 - 25 of 32

Show periods at a time.

▲ Label	Earliest start	Latest stop
Antiquité galo-romaine	1	499
Antiquité tardive	300	499
Dryas récent	-11000	-9600
Haut Empire	1	299

Moyen Âge

Permalink: <http://n2t.net/ark:/99152/p0rjd9qx9>

Original label Moyen Âge (fra-latn)
Alternate labels Middle Ages (eng-latn)
Spatial coverage dec... France
Spatial coverage France
Start 476
 (ISO value: 476)
Stop 1492
 (ISO value: 1492)
Notes in source by Isabelle Catteddu
Editorial notes http://www.inrap.fr/archeologie-preventive/Ressources/Dossiers-multimedias/Chronologie/Chronologie-des-periodes-de-l-histoire-et-de-l-archeologie/p-12506-Periodes-version-texte-htm?periode_id=10

Moyen Âge classique	1000	1299
Mésolithique	-9600	-5500
Mésolithique ancien	-9600	-8030
Mésolithique moyen	-8030	-6900
Mésolithique récente et finale	-6900	-5500
Néolithique	-6000	-2100
Paléolithique inférieur	-800000	-300000
Paléolithique moyen	-300000	-40000
Paléolithique supérieur	-40000	-12500
Période contemporaine	1800	1999
Période moderne	1492	1789
Renaissance	1466	1699
Siècle des Lumières	1600	1799
Siècles des l'Industrie	1800	1999
bas Moyen Âge	1300	1499
haut Moyen Âge	400	999
l'interstade de Bølling/Allerød	-12500	-11000

The purpose of using Perio.do in ARIADNEplus is different to the use of AAT data enrichment - rather than aligning data to common identifiers, Perio.do will be utilized in a subsequent data enrichment stage to enrich records already indexed with named periods (such as "Bronze Age", "Iron Age" etc.) where the dates associated with these periods are not already made explicit in the input data - providing start/end

dates to make these records comparable and searchable in the same way as other date information. If your data already contains suitable start/end dates then you need do nothing more on this. If your data contains named period terms without any further explanation then read on.

Establish the reference authority in use for the named periods present in your data. This should be an authoritative source - such as an existing published list of period names with associated dates, or a book or other citable publication exhibiting usage of the period terms and describing the dates associated with them. A suitable reference authority may already exist within Perio.do, so the first recommended course of action would be to examine the 'canonical' period list within Perio.do (i.e. the full set of curated period records) for an existing authority list that your data might agree with. If an appropriate entry is found you will need to identify the URI (identifier) of the authority list (the "*permalink*" as mentioned previously). To introduce a slight complication there are (currently) two versions of the Perio.do client available - "*Period Collection*" in the older client application (hosted at <https://test.perio.do/>) is synonymous with "*Authority*" in the newer client application (hosted at <https://client.perio.do/>). If using the older client application then select the "*Collection*" tab and click the title of the required "*Period Collection*" – the permalink URI is then displayed in the metadata of the selected collection. If using the newer client application then select the required "*Authority*" to filter the canonical list to terms from that authority only – selecting any of the resultant list of terms will then display the Authority permalink (among other data fields). As an example, the URI of the aggregated list of named periods as used in the first ARIADNE project is <http://n2t.net/ark:/99152/p0qhb66>. Once the URI has been identified, it will need to be communicated to the data ingestion/aggregation team, as it will be utilized in a subsequent data enrichment stage.

If no suitable authority list for your data is currently present in the Perio.do canonical list then a new list may be created and submitted. Contributing new period information to the Perio.do gazetteer is achieved via the Perio.do client application. If you do not already have one you will need to first set up an [ORCID](#) account to identify yourself on Perio.do. The [online user guide](#) provides detailed instructions on how to [create or edit your own](#) set of named periods, and how to [contribute your data](#) to the canonical period list. The new data is submitted to Perio.do using the client application in the form of a 'patch' that will need to be approved by Perio.do administrators prior to merging it into the canonical dataset, so there will be a delay from submission until your period data becomes visible in Perio.do. Once the data is visible you will be able to identify the Period Collection / Authority permalink (URI) that uniquely identifies your data, as described previously.

7.3. WGS84

The World Geodetic System 1984 (WGS84, aka EPSG:4326) is a standard geographic coordinate system (GCS), comprising a global horizontal and vertical datum and a coordinate system used (particularly by GPS systems) to express global positioning on the surface of the Earth. For ARIADNEplus purposes all spatial coordinates are to be expressed using WGS84. Other GCS also exist so coordinates in local datasets may require normalisation/transformation prior to aggregation. Use

of a common GCS to express spatial information will improve opportunities for cross searching the integrated datasets.

WGS84 coordinates are expressed in degrees – the horizontal position is the longitude (having values between -180° and $+180^{\circ}$, relative to a fixed datum of 0° at Greenwich, UK) and the vertical position is the latitude (having values between -90° and $+90^{\circ}$, relative to a fixed datum of 0° at the equator). As an example, the global location of the Leaning Tower of Pisa (Torre di Pisa) is given by WGS84 coordinates latitude 43.723056 ($43^{\circ} 43' 23''$ N), and longitude 10.396389 ($10^{\circ} 23' 47''$ E).

WGS84 coordinates may be serialized in various machine readable formats. In terms of the AO-Cat ontological model the coordinate values will be properties of class `ao:SpatialRegion` and its associated sub-classes (`ao:SpatialRegionPoint`, `ao:SpatialRegionBBox`, `ao:SpatialRegionPolygon`). The AO-Cat properties `has_latitude` and `has_longitude` are specified with a range data type of `xsd:decimal` – so the example (in Turtle RDF format) would be expressed as:

```
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix ao: <http://ariadne-infrastructure.eu/ns/> .
@prefix : <http://ariadne-infrastructure.eu/data/> . #this example
only

:x a ao:SpatialRegionPoint ;
  rdfs:label "Location of the Leaning Tower of Pisa"@en ;
  ao:has_latitude 43.723056 ;
  ao:has_longitude 10.396389 .
```

Note the importance of specifying the precision of the spatial data provided. See [Section 3.2.1](#).

8. Aggregation Pipeline Summary

In summary, the pipeline proceeds through the following stages:

1. Get the go-ahead to proceed, according to the workplan. If in doubt contact the WP4 lead (Julian Richards at UoY_ADS)
2. Confirm the collections to be aggregated and confirm which ARIADNE fundamental data category they belong to (see Section 3.4)
3. Use Activity Dash to keep track of the workflow pipeline or to communicate with the aggregation task force (Section 4)
4. Confirm the mechanism for data supply and frequency of update (Section 5)
5. Use the Timespans tool to normalise the dates in your records, if needed (Section 5.3)
6. Map the data to the AO-Cat using the 3M mapping tool (Section 6)
7. Create AAT subject mappings (Section 7.1)
8. Create Period0 period definitions (section 7.2)
9. Check spatial coordinates comply to WGS84 (Section 7.3)
10. Check that the data displays OK in the staging portal:
<https://ariadne-portal-staging.d4science.org/>
11. Give the green light to publish records on the ARIADNEplus public portal:
<http://portal.ariadne-infrastructure.eu>
12. Sign it off and have a well-deserved break!

All support is managed either via Activity Dash or via the ARIADNEplus [helpdesk](#) in D4Science. (See https://ariadne.d4science.org/group/ariadneplus_project and click on the tab "HelpDesk. Queries will be allocated to the appropriate specialist according to the category of ticket selected).

9. Appendix 1: Properties of AO_Data_Resource

Domain	Property	Scope note	range	Optional/ mandatory	one/ many
AO_Enity	has_identifier	an identifier for the resource in some namespace other than the provider's namespace	xsd:string	Optional	one
	has_type	a term from a controlled vocabulary	AO_Concept	Mandatory	many
	has_title	one title identifying the resource in a specific language. This	xsd:string	Mandatory	one per language
	has_description	a free text description of the resource	xsd:string	Optional	one
AO_Data_Resource	collection or individual item?	the nature of the resource, whether it's a collection with multiple items or an individual data resource	yes/no	Mandatory	one
	has_original_id	the local identifier of the resource supplied by the content provider	xsd:string	Mandatory	one
	has_ARIADNE_subject	one the fundamental archaeological categories defined by ARIADNE	AO_Concept	Mandatory	many
	has_native_subject	an original subject in the providing institution	AO_Concept	Mandatory	many
	has_language	the language(s) of the resource	AO_Concept (from lexvo.org)	Mandatory	many
	was_created_on	the date when the AO_Data_Resource was first made available online by the Publisher	xsd:dateTime	Mandatory	one
	has_landing_page	the original landing page	rdfs:Resource	Optional	one
	has_access_policy	the statement of policy (typically, on an organization's website) for the data resource	rdfs:Resource	Optional	one
	has_access_rights	information about who can access a data resource the resource or an indication of the data resource's security status	xsd:string	Mandatory	many
	has_extent	the size of the AO_Data_Resource (i.e., number of members in a collection, number of records in a dataset, etc.)	xsd:string	Optional	one
	has_temporal_coverage	the temporal region covered by the content of the resource	AO_Temporal_Region	Optional	one

	has_spatial_coverage	the spatial region covered by the content of the resource	AO_Spatial_Region	Mandatory	one
	has_spatial_precision	the precision with which the region is specified. The value is an instance of class AO_Dimension and can be given using two properties: has_value: a literal giving the amount of precision has_unit: a literal giving the unit of measure of the precision	AO_Dimension	Optional	one
	refers_to	the AO_Entity(ies) the resource refers to	AO_Entity	Mandatory	many
AO_Resource	was_issued	the date of formal issuance (e.g., publication) of the resource by the publisher	xsd:dateTime	Mandatory	one
	was_modified	the most recent date on which the resource was modified by the publisher	xsd:dateTime	Mandatory	one
	has_publisher	an agent responsible for making the resource publicly accessible	AO_Agent	Mandatory	one
	has_contributor	the agent primarily responsible for contributing the description of the resource to the ARIADNE Content Cloud	AO_Agent	Mandatory	one
	has_creator	the agent primarily responsible for creating the resource	AO_Agent	Mandatory	one
	has_owner	the legal owner of the resource, who holds the legal responsibility	AO_Agent	Mandatory	one
	has_responsible	the agent holding the scientific responsibility of the resource	AO_Agent	Mandatory	one

Agent

Domain	Property	Scope note	range	Optional/ mandatory	one/ many
Agent	An agent can be a person or a group. It has the following properties				
	has_name	the name of the agent	xsd:string	Mandatory	one
	has_agent_identifier	the identifier of the agent	xsd:string	Optional	many
	has_email	the email of the agent	xsd:string	Desirable	many
	has_homepage	the homepage of the agent	rdfs:Resource	Optional	one

In addition, if it is a person it also as a mandatory institution

has_institution	the institution where a person agent belongs	xsd:string	Optional	one
---------------------------------	--	------------	----------	-----

Spatial Region

Domain	Property	Scope note	range	Optional/ mandatory	one/ many
AO_Spatial_Region	every spatial region has the following properties				
	has_place_name	the name of the region	xsd:string	Mandatory	many
	has_coordinate_system	the coordinate system	xsd:string	Optional	one
	has_country_code	ISO 3166	rdfs:resource	Mandatory	One
	and is one of:				
	a point on the surface of the Earth identified by latitude and longitude;				
	has_latitude	latitude	xsd:decimal	Mandatory	one
	has_longitude	longitude	xsd:decimal	Mandatory	one
	a polygon as represented by GIS systems;				
	has_polygonal_representation	polygon as an XML data structure	rdfs:XMLLiteral	Mandatory	one
	a rectangular region on the surface of the Earth identified by its four vertices;				
	has_bounding_box_min_lat	minimal latitude	xsd:decimal	Mandatory	one
	has_bounding_box_min_lon	minimal longitude	xsd:decimal	Mandatory	one
	has_bounding_box_max_lat	maximal latitude	xsd:decimal	Mandatory	one
	has_bounding_box_max_lon	maximal longitude	xsd:decimal	Mandatory	one
	any region identified by a IRI in a standard gazetteer, such as Geonames for modern places, Pleiades for ancient places.				

has_place_IRI	the IRI of the place	rdfs:Resource	Mandatory	one
It may also be useful to record the degree of spatial precision				
has_spatial_precision	the precision with which the region is specified	AO_Dimension	Optional	one
has_value		rdfs:XMLLiteral	Optional	one
has_unit		rdfs:XMLLiteral	Optional	one

Temporal Region

Domain	Property	Scope note	range	Optional/mandatory	one/many
AO_Temporal_Region	A temporal region is one of: a temporal period defined in periodO and identified by a IRI				
	has_period	IRI of the periodo in periodO	rdfs:Resource	Mandatory	one
	a concept representing a period in some local vocabulary of the provider				
	has_native_period	concept	AO_Concept	Mandatory	one
	An interval defined by a beginning and an ending time point, both xsd:dateTime				
	from	starting time point of the interval	xsd:gYear	Mandatory	one
	until	ending time point of the interval	xsd:gYear	Mandatory	one