

ARIADNE Catalogue: Entering Data Resources A short user guide

Document history

Version	Content	Author	Date
0.1	Initial draft, based on discussions at ARIADNE	J.D. Richards	Apr. 3, 2015
	technical meetings in Leiden and Siena, and upon ARIADNE ACDM Version 2.41		
0.2	Revised draft based on comments by Carlo Megini, CNR	J.D. Richards & C.Megini	Apr. 9, 2015
0.3	Revised following review by Ulf Jakobsson	J.D. Richards & U.Jakobsson	Apr. 21, 2015
0.4	Alteration from dct:creator to dct:contributor	J.D. Richards & J.Fihn	May 4, 2015
0.5	Clarification that a single ACDM ArchaeologicalResource should have only one class	J.D. Richards & Dimitra Nefeli	Sept 1, 2015
0.6	Add ACDM samples	Dimitra Nefeli & J.D, Richards	Sept 17, 2015
1.1	Add workflow section	J.D. Richards	Sept 22, 2015
1.2	Changes to Workflow section and ACDM samples, adding table of contents	Dimitra Nefeli	Sept 30, 2015

Table of Contents

Purpose	Error! Bookmark not defined.		
1.0 Introduction	3		
2.0 The workflow	3		
Stage 1	3		
Stage 2a	3		
Stage 2b	3		
Stage 3	3		
Stage 4	4		
Stage 5:	4		
3.0 ACDM key fields	4		
ArchaeologicalResource	4		
DataResource	6		
i) Collection	6		
ii) Database	6		
iii) Dataset	6		
iv) GIS	6		
ArchaeologicalResourceType	6		
TemporalRegion	8		
SpatialRegion	9		
4.0 ACDM samples	9		
Collection level record	9		
MAPPING archiveSchema to ACDM	9		
Example of individual record for one textual document	12		
MAPPING extended DC (GLL) to ACDM	12		

Purpose

This document is intended as a gloss to the formal specification of the ACDM, and must always be used in conjunction with it. The ACDM describes the data model underlying the catalogue developed by the ARIADNE project for describing the archaeological resources that are made available by the partners of the project for the purposes of discovery, access and integration. These resources include: data resources, such as datasets and collections; services; and language resources, such as metadata formats, vocabularies and mappings.

This guide covers data resources only and is intended for use by data providers as an aid to the correct registration of their data resources in the ARIADNE registry. It is addressed to cultural institutions, private or public, who wish to describe their data resources in order to make them known to e-infrastructures.

1. Introduction

Data resources represent the various digital resources owned by the Ariadne partners and lent to the project for discovery, access and possibly integration. Data resources are categorized as collections or datasets, and have also been further divided into databases and GIS, as these are special categories of resources already available that users may wish to locate. As data resources of new types are introduced into the Catalogue, the model will be extended to accommodate them.

Data providers may upload their data to the registry via semi-automated harvesting of Excel spreadsheets, or OAI-PMH services, or they may enter individual resources manually, using the MORE tool. In each case the correct mapping of their metadata fields to the ACDM is essential to ensure that the ARIADNE resource discovery service will be usable. The purpose of this short guide is to provide guidance to partners on the data upload workflow and the mapping of key fields to the appropriate fields in the ACDM.

2. The workflow

Stage 1: Following agreed data upload schedule, contact DCU to agree best method for upload, e.g. OAI-PMH harvest, Excel spreadsheet, or manual entry to records in the Registry. (**Contact point**: Dimitris Gavrilis or Dimitra-Nefeli Makri) If OAI target is planned then use Section 3.0 below to ensure a correct mapping.

Stage 2a: Ensure that consistent period terms are provided for the region/country covered by the data, and ensure that these have been listed in the spreadsheet with start and end dates for PERIODO (**Contact point**: Lisa Fentress)

Stage 2b: Ensure that consistent subject terms can be provided for each entry. These must either conform to the Getty Art & Architecture thesaurus or partners must use the USW mapping tool to map their native terms to the AAT (**Contact point**: Doug Tudhope)

Stage 3: Upload your data! This can be checked in MORE (you can find guidelines in the following link: http://support.ariadne-infrastructure.eu/files/MORe_guidelines.pdf) and in the prototype portal interface: http://ariadne-portal.dcu.gr/laravel/.

Stage 4: Once partners are happy that their data is mapped correctly to the ACDM and is displaying correctly they need to supply information to enable any derived fields to be completed. They should use MORE to associate any period terms with the PeriodO permanent identifier. The partners will be able to enrich their data using the corresponding enrichment service. The result will be an enriched record, in which temporal information will be enhanced by the relevant terms of PeriodO.

Stage 5:_Finally, DCU will populate the "derived data fields" for subject and period, using the information supplied at Stage 2a and 2b

3. ACDM key fields

ArchaeologicalResource

All entries in the registry are instances of *ArchaeologicalResource*. This guide is concerned only with the subclass of DataResource.

The following table gives the attributes of ArchaeologicalResource in general, which must also be provided for subclass DataResource:

Attribute	Usage Notes	Obligation M(andatory) O(ptional)	UI control	Example
dct:title	As far as possible this should be a useful term, such as the name of a archaeological site	M	input text	Stonehenge
dct:description	This is a free text description of the resource	0	textarea	
dct:issued	Date when the record was first uploaded to the registry	М	date picker	2013-04-24
dct:modified	Most recent date on which the resource was changed, updated or modified.	M	date picker	2013-04-24
originalId	The original identifier of the resource supplied by the content provider	M	input text	This should be the ID of the resource in the catalogue of the content provider e.g. SND 2004-001
dct:identifier	The ARIADNE identifier that will be computed from the original. This	M (automatic)	input text	

dcat:keyword	could look like: www.ariadne.eu + prefix of the provider + ID of the resource	0	input text	
dct:language	If the resource is multilingual, more than one value must be given for this property	M	input text (autocompl ete)	
dcat:landingPage		0	input text	Where the resource is available online this should, as far as possible, be a persistent identifier e.g. http://dx.doi.org/10.5284/1021540
accessPolicy	URI to statement of policy (typically, on an organization's website)	0	input text	e.g. ADS Terms and Conditions
dct:accessRights	A statement of any access restrictions on the data resource	М		e.g. Creative Commons 2.0 Licence
dct:rights	Information about rights held in and over the data resource	0		
dct:audience	The primary audience(s) of the resource	0		

All archaeological resources, and therefore all data resources, also have the following associations:

dct:publisher: associates any archaeological resource with an agent responsible for making the resource publicly accessible (via download, or API, or other). For example, the ADS has made the English Heritage Excavation Index available online.

dct:contributor: associates any archaeological resource with an agent primarily responsible for describing the resource in the Catalogue. This may be the same as the publisher, where ADS has uploaded the Excavation Index to the Catalogue.

owner: associates any archaeological resource with an agent that is the legal owner of the resource. For example, in this case English Heritage would be the owner of the resource.

The class **agent** is used to register information about organisations that may fall into any of the above three categories. Named individuals should be listed as legalResponsible, scientificResponsible, or technicalResponsible.

NOTE!

Any one ArchaeologicalResource should contain metadata for only one class of resource such as collection, dataset, database, service, vocabulary etc. Where there are multiple classes of resource then a separate ArchaeologicalResource record should be created for each so that they can have

separate identifiers. For example, if a provider has one collection containing 10.000 reports, the following records should be created:

- 1. One record with the metadata of the collection, using its identifier (mandatory element);
- 2. 10,000 records with the metadata of each textual document, which will be linked to the collection using the collection's identifier (isPartOf relation).

DataResource

DataResources will also be categorized as being one of 4 types:

i) Collection

A collection as an aggregation of resources, called the items in the collection. Being aggregations, collections are akin to datasets, but with the following, important difference: the items in a dataset are data records of the same structure (see definition of Dataset below). In contrast, the items in a collection are individual objects different from records (e.g., images, texts, videos, etc.) or are themselves data resources such as collections, datasets, databases or GIS; for instance, a collection may include a textual document, a set of images, one or more datasets and other collections. Thus, the ArchaeologicalResourcetype of 'Fieldwork archive' will generally be a Collection.

For convenience, however, collection items that are textual documents are explicitly represented in the Catalogue as instances of the class TextualDocuments and associated to the collection where they belong via the association dct:isPartOf. Thus the ADS grey literature library is a Collection, but individual reports also appear at item level in the catalogue as TextualDocuments that belong in that collection.

ii) Database

Databases are defined as a set of homogeneously structured records managed through a Database Management System, such as MySQL. For example, the ArchaeologicalResourcetype of Burial databases will generally be within this subclass. There may be databases within Collections, which may also be entered as individual items in the Catalogue.

iii) Dataset

A dataset is defined as a set of homogeneously structured data records, consisting of fields carrying data values, that are not managed through a Database Management System or a GIS. Excel spreadsheets may be categorised as Datasets, for example.

iv) GIS

A resource held in a GIS will be categorised as a GIS.

In addition, all Data Resources will have the following attributes:

ArchaeologicalResourceType associates any data resource with one or more categories (types), drawn from the following list:

- 1. Fieldwork archives
- 2. Event/intervention resources, such as grey literature reports

- 3. Sites and monuments databases or inventories
- 4. Scientific datasets, such as databases of radiocarbon dates
- 5. Artefact databases, or artefact image collections
- 6. Burial databases
- 7. Other

This is not a formal or comprehensive categorisation, and in due course it will be extended to define other key groupings. However it was derived from a preliminary survey of data resources held by partners to be made available via ARIADNE, as an aid to identifying candidates for potential integration. It is a mandatory field.

1. Fieldwork archives

These entries describe collections of one or more data files, generally of one or more types, that are derived from archaeological fieldwork. Such collections will typically comprise text files, spreadsheets, databases, images, and potentially CAD and GIS files. If the resource is purely a text file then it should be categorised as 2 below (an event/intervention/ grey literature resource). Normally a single metadata record will be provided in the registry for the collection as a whole.

2. Event/intervention resources, such as grey literature reports

This category of resource is closely related to sites and monuments inventories (3) but provides instead a record of where archaeological fieldwork has been carried out, as opposed to where known sites are. In database terms there will frequently be a one-to-many relationship between monuments and events, as each site will have been subject to multiple investigations over time. It is the events, however, that provide most information about a site or monument. A special category of event-based resource held by several partners are effectively bibliographic metadata records for unpublished fieldwork reports, or 'grey literature'. Normally a metadata record will be provided for each report, or record of an intervention.

3. Sites and monuments databases or inventories

Most European countries and regions hold inventories of known archaeological sites and monuments within their area. These are often developed for management purposes, but also provide an invaluable research resource. In addition several partners are also offering ARIADNE datasets of sites of specific classes or periods. Although not comprehensive they can usefully be searched alongside the management-driven records.

Such inventories are generally characterised by fields for place, period and multiple keywords for site or monument type. Given that most archaeological research questions are cross-border (e.g. "Where are all the Bronze Age sites in southern Europe?") there is a value-added from combining inventories that ARIADNE is uniquely placed to provide. Therefore partners may choose to upload them to the registry at item level, with one record per site.

4. Scientific resources

These include databases of radiocarbon dates, environmental or faunal datasets, or the results of scientific analyses. At the present time a single registry entry will be sufficient to describe each such resource.

5. Artefact databases, or image collections

Several partners have dedicated artefact databases, but they may also be present in fieldwork resources. Some databases may include images of finds, or thumbnail links to separate image files. Generally only a single collection-level record will be uploaded to the registry.

Databases of coins have already been singled out for experimentation in item level integration, but would just feature as a single record in the registry.

6. Burial databases

These are a special class of database where the individual grave represents the item level. The attributes recorded may include finds as well as physical anthropological information. Cemeteries may also appear as single records in sites and monuments and events databases, which could also provide a pointer to the richer item level datasets. Databases of burials have already been singled out for experimentation in item level integration, but would just feature as a single record in the registry.

7. Other

Any other category that does not fit into those above.

In addition to the above, Data Resources will also be described by the following elements in order to allow What, When and Where searching:

dct:subject associates any archaeological resource with a controlled subject term drawn from the Getty AAT vocabulary. A resource must have at least one such subject. This is further categorized as (so any provider should only give one of the following):

- **provided-subject** associates any archaeological resource with zero, one, or more manually specified subjects drawn from the Getty AAT vocabulary;
- **derived-subject** associates any archaeological resource with zero, one, or more subjects, automatically derived from mapping native vocabularies to the Getty AAT vocabulary.

native-subject associates any archaeological resource with a subject from a controlled vocabulary in use by the original owner of the resource. It could, for example, be a subject term drawn from the Historic England Thesaurus of Monument types, or from the Thesaurus of Archaeological object types.

native-vocabulary is the URI of the vocabulary where the native subject(s) of an archaeological resource are drawn from. The underlying assumption here is that there is at most one such vocabulary.

TemporalRegion

All data resources must be described by at least one term that indicates the archaeological time period to which they relate. This may be either a start and end date, or a named period (e.g.,

neolithic). In the former case, the extremes of the interval are given as vaues of the *from* and *to* attributes. In the latter case, the named period is given as value of the *periodName* attribute. In this case it is also necessary to draw the period term from a specified controlled vocabulary for which start and end dates are supplied.

SpatialRegion

All data resources must have at least one spatial region defined, in one of four forms:

- (i) a region identified by latitude and longitude expressed via the *lat* and *lon* attributes, respectively;
- (ii) a bounding box identified by four vertices (expressed via the boundingBoxMinLat, boundingBoxMinLon, boundingBoxMaxLat and boundingBoxMaxLon attributes);
- (iii) a postal address (expressed via the address, numberInRoad, postcode and country attributes); or
- (iv) a named place (expressed via the *placeLabel* attribute).

4.0 ACDM samples

Let's say that a provider has a collection containing some thousands of textualDocuments. The collection is encoded in a schema called "archiveSchema" while the textualDocuments are encoded in a customized Dublin Core. You can find below the native and the ACDM records, as well as the mapping between them.

Collection level recordMAPPING archiveSchema to ACDM

<acdm;nativeSubject> <skos:Concept>

<skos:prefLabel> archive.Coverage.subjects.subject

<acdm:publisher>

<foaf:name> Archaeology Data Service

<acdm:typeOfAnAgent> Organization

<acdm:contributor>

<foaf:name> archive.Actors.actor[type="creator"].name OR

archive.Actors.actor[type="creator"].organisation

<acdm:typeOfAnAgent> Person / organization

<dcterms:title> archive.title

<dc:description> archive.description

<dcterms:issued> archive.published.firstReleased
<dcterms:modified> archive.published.lastUpdated

<acdm:originalId> archive.id

<dc:language> er

<dcat:landingPage > "http://dx.doi.org/" + archive.doi

<acdm:accessPolicy>

"http://archaeologydataservice.ac.uk/advice/termsOfUseAndAccess"

<acdm:archaeologicalResourceType> archive.acdmType

<acdm:temporal>

Natve.xml

The aim of this resource is to make available unpublished fieldwork reports in an easily retrievable fashion. The number of reports available is increasing steadily through the OASIS project in England and Scotland. These unpublished reports cover archaeological interventions such as: watching briefs, excavation reports, building surveys, desk-based assessments and specialist analysis.

```
</description>
<version>1</version>
<types>
      <type>Text Collection</type>
</types>
<acdmType>Event/Intervention resources</acdmType>
<language>eng</language>
<created>
      <from>2005-05-01</from>
      <to>2015-06-12</to>
</created>
<published>
      <firstReleased>2005</firstReleased>
</published>
<Actors>
      <actor type="funder">
             <organisation>Historic England/organisation>
      </actor>
</Actors>
<Coverage>
      <subjects>
             <subject type="LCSH">Archaeology</subject>
      </subjects>
      <coordinates>
             <box>
                    <northLatitude>60.802064</northLatitude>
                    <eastLongitude>1.97241120</eastLongitude>
                    <southLatitude>49.781264</southLatitude>
                    <westLongitude>-10.991455</westLongitude>
             </boundingBox>
      </coordinates>
```

```
<locations>
                     <location type="British Isles country">England</location>
                     <location type="British Isles country">Wales</location>
                     <location type="British Isles country">Scotland</location>
                     <location type="TGN">World, Europe, United Kingdom [7008591]/location>
             </locations>
      </Coverage>
      <repository>Archaeology Data Service</repository>
      licence type="other">Terms and conditions apply to reuse.
</archive>
```

ACDM.xml

<acdm:ariadne xmlns:acdm="http://registry.ariadne-

infrastructure.eu/" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dbpedia-

owl="http://dbpedia.org/ontology/" xmlns:foaf="http://xmlns.com/foaf/" xmlns:dcmitype="http://purl.o rg/dc/dcmitype/"xmlns:vcard="http://www.w3.org/2006/vcard/ns#" xmlns:dcterms="http://purl.org/dc /terms/" xmlns:oai dc="http://www.openarchives.org/OAI/2.0/oai dc/" xmlns:rdfs="http://www.w3.or a/2000/01/rdf-

schema#" xmlns:dcat="http://www.w3.org/ns/dcat#"xmlns:aat="http://vocab.getty.edu/aat/" xmlns:o wl="http://www.w3.org/2002/07/owl#" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntaxns#" xmlns:sch="http://purl.oclc.org/dsdl/schematron" xmlns:skos="http://www.w3.org/2004/02/sko s/core#"xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">

<acdm:ariadneArchaeologicalResource>

```
<acdm:collection>
      <acdm:publisher>
             <foaf:name>Archaeology Data Service</foaf:name>
             <acdm:typeOfAnAgent>Organization</acdm:typeOfAnAgent>
      </acdm:publisher>
      <acdm:nativeSubject>
             <skos:Concept>
                    <skos:prefLabel>Archaeology</skos:prefLabel>
             </skos:Concept>
      </acdm:nativeSubject>
      <dcterms:title>Library of Unpublished Fieldwork Reports</dcterms:title>
      <dc:description>
```

The aim of this resource is to make available unpublished fieldwork reports in an easily retrievable fashion. The number of reports available is increasing steadily through the OASIS project in England and Scotland. These unpublished reports cover archaeological interventions such as: watching briefs, excavation reports, building surveys, desk-based assessments and specialist analysis.

```
</dc:description>
<dcterms:issued>2005</dcterms:issued>
<acdm:originalId>1093</acdm:originalId>
<dc:language>en</dc:language>
<acdm:accessPolicy>
http://archaeologydataservice.ac.uk/advice/termsOfUseAndAccess
</acdm:accessPolicy>
<dcterms:accessRights>ADS Terms and
```

Conditions</dcterms:accessRights> <acdm:spatial> <acdm:placeName>World, Europe, United Kingdom

[7008591]</acdm:placeName>

<acdm:boundingBoxMinLat>49.781264</acdm:boundingBoxMinLat>

<acdm:boundingBoxMinLon>-

10.991455</acdm:boundingBoxMinLon>

<acdm:boundingBoxMaxLat>60.802064

</acdm:boundingBoxMaxLat>

<acdm:boundingBoxMaxLon>1.97241120

</acdm:boundingBoxMaxLon>

</acdm:spatial>

<acdm:archaeologicalResourceType>Event/Intervention

resources</acdm:archaeologicalResourceType>

<dc:identifier>

http://registry.ariadne-infrastructure.eu/collection/11972121

</dc:identifier>

</acdm:collection>

</acdm:ariadneArchaeologicalResource>

</acdm:ariadne>

Example of individual record for one textual documentMAPPING extended DC (GLL) to ACDM

<acdm:nativeSubject> <skos:Concept>

<skos:prefLabel> dc:subjectMonuments, dc:subjectFinds

<dcterms:title> dc:title <dcterms:issued> dc:published <acdm:originalld> dc:identifier

<dc:language> en

<dcat:landingPage> "http://dx.doi.org/" + dc:identifier

<acdm:accessPolicy>http://archaeologydataservice.ac.uk/advice/termsOfUseAndAccess

<dcterms:accessRights> ADS Terms & Conditions

<acdm:archaeologicalResourceType> Event/intervention resources

Natve.xml

<oai_dc:dc xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:gll="http://archaeologydataservice.ac. uk/archives/view/greylit/" xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/" xmlns:xsi=" http://www.w3.org/2001/XMLSchema-

instance"xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/

http://www.openarchives.org/OAI/2.0/oai_dc.xsd">

<dc:title>Land at Dunkeswell Airfield Dunkeswell Devon</dc:title>

<gll:author>Sommers, M.</gll:author>

<gll:otherBib>SCCAS Report Number 2004/88/gll:otherBib>

<dc:description>3 sheets of A4 paper stapled togeather.</dc:description>

<dc:publisher>Cotswold Archaeology</dc:publisher>

<gll:published>2004</gll:published>

<dc:date>2012-08-08 16:52:16.0</dc:date>

<dc:format>1 pdf</dc:format>

<dc:identifier>10.5284/1000544</dc:identifier>

<gll:link>

http://archaeologydataservice.ac.uk/archives/view/greylit/details.cfm?id=146

```
</gll:link>
  <gll:source>OASIS</gll:source>
  <dc:coverage>ENGLAND, SUFFOLK, ST EDMUNDSBURY, CLARE</dc:coverage>
  <gll:place>Ipswich</gll:place>
  <gll:site>CLA 041 16 Nethergate Street, Clare</gll:site>
  <dc:coverage type="spatial">TL 7677 4511</dc:coverage>
</oai_dc:dc>
```

ACDM.xml

<acdm:ariadne xmlns:acdm="http://registry.ariadne-

infrastructure.eu/" xmlns="http://archaeologydataservice.ac.uk/advice/archiveSchema" xmlns:dc=" http://purl.org/dc/elements/1.1/" xmlns:dbpedia-

owl="http://dbpedia.org/ontology/" xmlns:foaf="http://xmlns.com/foaf/"xmlns:dcmitype="http://purl.org/dc/dcmitype/" xmlns:vcard="http://www.w3.org/2006/vcard/ns#" xmlns:dcterms="http://purl.org/dc/terms/" xmlns:rdfs="http://www.w3.org/2000/01/rdf-

schema#" xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"xmlns:dcat="http://www.w3.org/ns/dcat#" xmlns:aat="http://vocab.getty.edu/aat/" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-

ns#" xmlns:sch="http://purl.oclc.org/dsdl/schematron"xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:gll="http://archaeologydataservice.ac.uk/archives/view/greylit/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">

<acdm:ariadneArchaeologicalResource>

<acdm:textualDocument>

<dcterms:title>Land at Dunkeswell Airfield Dunkeswell

Devon</dcterms:title>

<acdm:originalId>10.5284/1000544</acdm:originalId>

<dc:language>en</dc:language>

<dcat:landingPage>http://dx.doi.org/10.5284/1000544</dcat:landingPage>

<acdm:accessPolicv>

http://archaeologydataservice.ac.uk/advice/termsOfUseAndAccess

</acdm:accessPolicy>

<dcterms:accessRights>ADS Terms and

Conditions</dcterms:accessRights>

<acdm:archaeologicalResourceType>Event/intervention

resources</acdm:archaeologicalResourceType>

< dcterms:isPartOf xmlns="">

http://registry.ariadne-infrastructure.eu/collection/10138474

</ dcterms:isPartOf > <dc:identifier xmlns="">

http://registry.ariadne-infrastructure.eu/textualDocument/10398968

</dc:identifier>

</acdm:textualDocument>

</acdm:ariadneArchaeologicalResource>

</acdm:ariadne>