DELIVERABLE

Project Acronym: CARARE
Grant Agreement number: 250445
Project Title: Connecting ARchaeology and ARchitecture in Europeana

D3.3.3 – Documented CARARE workflow

Revision: Final

Author: D. Gavrillis and C. Dallas, DCU
V. Tzrouvaras and K. Paradalis, NTUA

<table>
<thead>
<tr>
<th>Dissemination Level</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PU</td>
<td>Public</td>
</tr>
<tr>
<td>CO</td>
<td>Confidential, only for members of the consortium and the Commission Services</td>
</tr>
</tbody>
</table>
## Revision History

<table>
<thead>
<tr>
<th>Revision</th>
<th>Date</th>
<th>Author</th>
<th>Organisation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>v.0.1</td>
<td>01/10/11</td>
<td>D. Gavrillis &amp; C. Dallas</td>
<td>DCU</td>
<td>Outline</td>
</tr>
<tr>
<td>v.0.2</td>
<td>26/10/11</td>
<td>D. Gavrillis &amp; C. Dallas</td>
<td>DCU</td>
<td>First draft</td>
</tr>
<tr>
<td>v.0.3</td>
<td>30/10/11</td>
<td>Kostas V. Pardalis</td>
<td>NTUA</td>
<td>Additions</td>
</tr>
<tr>
<td>v.0.4</td>
<td>11/11/11</td>
<td>D. Gavrillis &amp; C. Dallas</td>
<td>DCU</td>
<td>Update</td>
</tr>
<tr>
<td>v.1</td>
<td>17/11/11</td>
<td>K. Fernie</td>
<td>MDR</td>
<td>Integration of final comments</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>from peer review</td>
</tr>
</tbody>
</table>

## Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.
Contents

1 EXECUTIVE SUMMARY 4

2 INTRODUCTION 5

2.1 Mapping and transformation of metadata 5

2.2 Quality assurance framework 5

3 AGGREGATOR’S WORKFLOW 6

3.1 Metadata upload 6

3.2 Mapping and Ingestion preparation workflow 7

3.3 Repository publishing workflow 8

3.3.1 Ingestion from MINT to MORE 9

3.3.2 Direct harvesting of CARARE compatible OAI-PMH repositories 10

3.4 Repository facilities workflow 11

3.5 Europeana delivery workflow 12

4 SUMMARY OF THE CARARE METADATA WORKFLOW 13

REFERENCES 14

ANNEX I. CARARE SCHEMA TO EDM SCHEMA MAPPING 15

4.1 CARARE Record example 21

4.2 EDM record example 29

ANNEX II. MAPPING AND IMPLEMENTATION GUIDELINES FOR THE CARARE METADATA SCHEMA 32
1 Executive Summary

This report documents the workflows present in the CARARE system and the flow of information from the content providers’ native repositories (in their native schemas) to Europeana (EDM).

Chapter 2 provides an introduction to the workflow.

Chapter 3 describes the Aggregation workflow, this is a complex process which requires that the metadata are exported, mapped, transformed, enriched and passed between a number of different systems: from the content provider’s native repository to MINT, the MORE repository and finally to Europeana. The stages in the workflow are:

- Metadata upload of the content provider’s native records to the MINT tool;
- Mapping and ingestion preparation workflow involves the process of mapping and converting native records to the CARARE schema;
- Repository publishing workflow offers providers two options:
  - preparing a SIP package to transfer information from MINT to MORE
  - direct harvesting of a CARARE schema compatible OAI-PMH repository by the MORE repository
- Repository facilities workflow describes the facilities available on the MORE repository to review data prior to delivery to Europeana
- Europeana delivery workflow consists of the steps required for CARARE records to be mapped to EDM and harvested by Europeana.

Chapter 4 provides a summary of the workflow.

Annex 1 contains the CARARE-EDM schema mapping together with an example record in CARARE schema format and its transformation to EDM.

Annex 2 provides mapping and implementation guidelines for the CARARE metadata schema.
2 Introduction

2.1 Mapping and transformation of metadata
Mapping is the main transformation that information is subjected to when entering and traversing the CARARE system. First, native records are mapped to the CARARE schema by the providers themselves or by using the MINT tool. This is necessary in order to convert all data to a common schema and thus, be able to perform other operations under this common reference. This operation produces the CARARE datastream.

Information can then be enriched in various ways, such as: convert spatial coordinates to a common reference, add semantic relations to other objects, add collection information identifiers, etc. This would produce the CARARE enriched datastream.

Finally, the enriched CARARE record is transformed to EDM, the schema that Europeana understands and can make use of. This transformation is realized by a set of complex rules that can be found in Annex I.

2.2 Quality assurance framework
The quality assurance framework that was designed and implemented for CARARE consists of a multi-stage process that ensures the quality of information harvested by Europeana. This quality assurance is summarized in the following:

Conceptual level:
- Mediating schema (CARARE) is as rich as possible, including all useful information that could be found in the content providers’ native schemata.
- The minimum number of elements that must be provided in each CARARE record (the mandatory elements) will lead to EDM records with adequate quality.

Technical level:
- The various tools MINT and MORE make all necessary structural checks (XML well-formedness validations)
- Mandatory elements checks
- Statistics monitoring of strongly recommended elements checks
- Logical checks (e.g. both X,Y coordinates are provided).
3 Aggregator’s workflow

The CARARE aggregation infrastructure consists of the MINT and MORE platforms, which are configured to ensure seamless mapping and ingestion of content provider metadata to CARARE. The overall system architecture is presented in figure 1 below.

Content providers can upload their native records to the MINT tool directly by using one of the metadata ingestion methods that are provided by the MINT ingestion platform (see below, section 3.1). The MINT tool provides tools for mapping and transforming the native metadata records received to the CARARE Schema, and providing them for ingestion to MORE using the SIP protocol.

Alternatively, content providers who are able to convert their native records directly to the CARARE schema are offered the option to expose them directly for harvesting/ingestion by MORE through OAI-PMH. MORE enriches the records, converts them to EDM and exposes them through OAI-PMH for harvesting by Europeana.

![Figure 1: CARARE system architecture](image)

3.1 Metadata upload

The metadata upload process involves the upload of the content providers’ native XML or CSV records to the MINT tool. This can be achieved through the following procedures:

- Using a remote FTP or HTTP upload of a single XML file. In this case the user is prompted to provide a valid URL using one of the two before mentioned protocols for remotely uploading metadata records either as a single XML file or as multiple files packaged in a ZIP archive.
• Using a direct HTTP upload of one XML file or a ZIP archive of a whole collection. In this case the user directly uploads the metadata records from his/her local computer. Using this option the user is also able to upload metadata records in CSV format, in that case, the user is also prompted to provide information regarding the field separator and the use of any special character as the “escape character”. For interoperability reasons it is recommended that users provide XML records.

• “FTP Upload” and “Server Filename” options are offered to support providers that do not have direct access to their native records over the internet or providers who have extremely large datasets which needs special handling by the MINT development team; these options are rarely used.

• OAI harvesting. An OAI-PMH V2 compatible harvester is implemented and exposed to the user as a metadata upload option for MINT. In this case the user is prompted to provide all the appropriate mandatory and optional parameters for the harvester, e.g. the base URL of the OAI-PMH repository, the namespace prefix, an optional set name and appropriate date values for filtering. The user is also able to validate the provided parameters and also fetch information that is provided by the OAI-PMH repository.

During the metadata upload process, the MINT tool analyzes the XML records and performs well-formedness and validation checks.

### 3.2 Mapping and Ingestion preparation workflow

The MINT mapping and ingestion preparation workflow process describes the process of converting native records to CARARE and packaging them for ingestion to MORE (see figure 2 below).

The content provider uploads their metadata to MINT using the methods that are presented in 2.1. The metadata are analyzed, validated and the native schema is inferred. As part of the schema inference procedure, the user is also prompted to select a structural element of the inferred metadata schema that will act as the wrapper of a unique metadata record.

Additionally, the user is given the option to select one of the elements to act as the main label of the records in order to provide better browsing and visualization of the uploaded records. This requires however, that the content provider use representative and complete records so that the full native schema is inferred. Next, the content provider is presented with the native schema along with the CARARE schema and proceeds to create the mapping between the two by establishing connections between elements in the native schema and corresponding elements in the CARARE schema.
The mapping tool constitutes the core functionality provided by the MINT platform. It offers the user the ability to map his/her native inferred metadata Schema to the CARARE Schema. The user is also supported by functionalities that allow browsing of the values of the various metadata elements, value mappings in cases where data normalization is needed and various functions for manipulating both the structural elements of the inferred Schema and the values these elements contain. The mappings that are created are validated by sampling the provided metadata records, instantiating CARARE records and validating them against the CARARE Schema. In this way users are able to quickly validate the resulting mappings and have an overview of the resulting metadata records that will be published to the MORE repository. This iterative process enables the user to fine tune the mappings and this also acts a quality assurance mechanism that is integrated on an architectural level to the MINT tool.

After the mapping is finalized, the content provider can publish his metadata to MORE. The publication step involves the actual transformation of the ingested metadata to the CARARE schema, their packaging to a format that MORE accepts (SIP packages that contain the native metadata, the CARARE metadata and the mapping definition itself) and finally the ingestion of the SIP package to MORE.

### 3.3 Repository publishing workflow

The MORE repository publishing workflow concerns the steps required for the ingestion of information. Two workflows have been implemented:

- Ingestion from MINT to MORE.
- Direct harvesting of content providers’ OAI-PMH repositories exporting records in CARARE format.
3.3.1 Ingestion from MINT to MORE.

Information is packaged using the SIP format which contains multiple records. This procedure goes beyond traditional approaches where one would send only the CARARE record. It stipulates that SIPs to be ingested provide the following:

1. Native metadata record
2. CARARE metadata record
3. Mapping mechanism used to move from native to CARARE
4. Administrative and technical metadata (e.g. user, content provider, timestamp, etc.).

This approach aims to address requirements related to recording provenance and digital preservation information. It provides for an auditing mechanism of the ingestion process, and allows content providers to examine, for each item, the native and CARARE records as well as administrative and technical metadata.

An important advantage of the approach followed is that it allows for version tracking within MORE (the CARARE repository). If an item has already been ingested into the repository, the system will ingest it as a newer version of the already ingested item (keeping its complete history). In the current version of the system, the user can browse the native, CARARE and EDM records of the latest version of each item.

![Figure 3: Mapping and versioning workflow](image)

The publishing workflow has been implemented using a REST service which accepts SIP packages. The workflow for publishing follows the steps below:
1. MINT creates the SIP package and triggers the service by supplying the URL of the SIP package.
2. SIP package is downloaded to MORE’s temporary space
3. SIP package is uncompressed and its structure validated. The package must contain an index (xml file containing contents). Content provider is recognized.

For each item:

4. Each item is validated (it must contain three datastreams: a) native record, b) CARARE record, c) XSLT mapping). In the case where the item has been harvested directly as CARARE, the latter two datastreams can be omitted.
5. All XML datastreams are validated.
6. Existing ingest based on the same item is located and if not, a new object is created on the repository.
7. Collection information is extracted and the collection registry is updated.
8. All datastreams are ingested into the repository.

If any errors occur during this process, an XML report is produced and returned to MINT through the associated web service.

![Figure 4: Publishing workflow](image)

### 3.3.2 Direct harvesting of CARARE compatible OAI-PMH repositories

A second workflow has been implemented for content providers who are exposing their content directly in the CARARE schema through OAI-PMH.
In this workflow:
1. MORE harvests the CARARE records exposed by the repository
2. Records are downloaded to MORE’s temporary space
3. The XML datastreams are validated.
4. Existing ingest based on the same item is located and if not, a new object is created on the repository.
5. Collection information is extracted and the collection registry is updated.
6. Datastreams are ingested into the repository
7. The transformations into EDM are made
8. Records are exposed them to the Europeana harvester

3.4 Repository facilities workflow
Once the content providers’ metadata are in the CARARE repository (MORE), they can perform specific operations before their records are published to Europeana. The workflow a user can follow inside the repository (shown in figure 5) is as follows:

1. User logs in (user is associated with a content provider)
2. View the overall report on the content provider he is associated with
3. User searches/browses for records
4. User locates a record and views its details
5. Inside the record view the user can
   a. Use the relation editor to add a relation to another object / manage relations
   b. Withdraw / publish the object to Europeana
   c. Enable / Disable the object (does not appear to other users searches)
   d. Display the completeness report on the object’s metadata (for CARARE schema only).
   e. Provide a quick report on the spatio-temporal information contained within the record.
   f. Label the record (labels are used like tags and aim at grouping objects together).
3.5 Europeana delivery workflow

The Europeana delivery workflow consists of the steps required for CARARE records to be mapped to EDM and harvested by Europeana.

The mapping function consists of producing Europeana Data Model (EDM) provided CHO out of items contained in CARARE records. Currently, provided CHO are produced for each Heritage Asset and Digital Resource in a CARARE record, following the mapping outlined in Annex 1 below.

The harvesting function is achieved through MORE’s OAI-PMH provider service (http://store.carare.eu:8080/oaprovier). Europeana’s harvester can trigger a harvest at any time. Currently no Sets are defined although this is technically possible.

A ‘publish’ flag is provided, and set by default, to allow content providers to exclude objects from Europeana publishing; only objects having their ‘publish’ flag set are exposed through OAI-PMH for harvesting. In the case that a content provider withdraws a set of already published records by resetting their ‘publish’ flag, these records will be withdrawn from Europeana’s servers after the next harvest (in current arrangements, scheduled once a month).
4 Summary of the CARARE metadata workflow

The full workflow of the CARARE system is documented in this report. It is supported by two interoperable applications, the MINT tool and the MORE repository, and is composed of four distinct components: the MINT mapping and ingestion preparation workflow, the MORE publishing workflow, the MORE repository functions workflow, and the Europeana harvesting workflow.

For easy reference, the complete CARARE metadata workflow can be summarised as follows:

Case 1: If content providers expose their content to the MINT tool (through OAI or via XML dumps).
  - The MINT tool harvests the content, makes the necessary transformations (converts them into CARARE records) and packages it using the SIP format (described in earlier reports – and thoroughly in the D2.5 Technical approach).
  - The CARARE repository harvests the SIP packages and ingests their content into the repository.
  - The CARARE repository makes the necessary transformations (converts the CARARE records to EDM) and exposes them to the Europeana harvester.

Case 2: If content providers expose their content directly into the CARARE schema through OAI.
  - The CARARE repository harvests the CARARE records, makes all necessary transformations (into EDM) and exposes them to the Europeana harvester.

In sum, the CARARE system architecture, composed of the MINT mapping and ingestion preparation tool, and the MORE monument repository, provides for an end-to-end lifecycle ensuring delivery of content providers’ metadata, originally in the form of native XML records, to the Europeana harvester, in the form of designated EDM provided CHOs. It also fully caters for necessary transformations so as to ensure successful harvesting by Europeana, and provides monitoring facilities to ensure metadata well-formedness and completeness.
References

CARARE Metadata Interoperability Tool: http://CARARE.image.ntua.gr/CARARE/

CARARE metadata schema outline 1.0: http://www.CARARE.eu/eng/Resources

CARARE Repository: http://store.CARARE.eu/

D2.2.3 - Metadata Mappings: http://www.CARARE.eu/eng/Resources

D2.2.5 – White paper on the CARARE technical approach: http://www.carare.eu/eng/Media/Files/White-paper-on-CARARE-technical-approach

D3.3.4 - Briefing paper on metadata mapping and the use of mapping tools: http://www.CARARE.eu/eng/Media/Files/D3.4-Briefing-paper-on-metadata-mapping-and-the-use-of-mapping-tools

D4.4.3 – Timetable and implementation plan

D4.4.4 - Report on the repositories established by each partner

An nex I. CARARE Schema to EDM Schema Mapping

In the CARARE schema the main elements is the “Heritage Asset” and the “Digital Resource”. These elements and their sub-elements have been mapped to the appropriate EDM elements as shown in the following table.

<table>
<thead>
<tr>
<th>EDM</th>
<th>CARARE</th>
<th>CARARE</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>edm:ProvidedCHO</td>
<td>Heritage Asset Identification/Record Information /ID</td>
<td>Digital Resource/Record information/ID</td>
<td>The value is entered in an rdf:about attribute</td>
</tr>
<tr>
<td>dc:contributor</td>
<td>Digital Resource/Actors/Name (when in contributor role)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dc:creator</td>
<td>Digital Resource/Actors/Name (when in creator role)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dc:date</td>
<td>Heritage Asset Identification/Characters/ Temporal/start date</td>
<td>Digital Resource/Temporal/Time span/start date</td>
<td></td>
</tr>
<tr>
<td>dc:date</td>
<td>Heritage Asset Identification/Characters/ Temporal/end date</td>
<td>Digital Resource/Temporal/Time span/end date</td>
<td></td>
</tr>
<tr>
<td>dc:date</td>
<td>Heritage Asset Identification/Characters/ Temporal/Display date</td>
<td>Digital Resource/Temporal/Display date</td>
<td></td>
</tr>
<tr>
<td>dc:date</td>
<td>Heritage Asset Identification/Characters/ Temporal/Scientific date</td>
<td>Digital Resource/Temporal/Scientific Date</td>
<td></td>
</tr>
<tr>
<td>dc:date</td>
<td>Heritage Asset Identification/Characters/ Craft/DateofLoss</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dc:description</td>
<td>Heritage Asset Identification/Description</td>
<td>Digital Resource/Description</td>
<td></td>
</tr>
<tr>
<td>dc:description</td>
<td>Heritage Asset Identification/Characters/ Craft/LastJourneyDetails/ MannerofLoss</td>
<td>Digital Resource/Note</td>
<td></td>
</tr>
<tr>
<td><strong>dc:description</strong></td>
<td>Heritage Asset Identification/Characters/Craft/LastJourneyDetails/Cargo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------------------------------------------</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td><strong>dc:format</strong></td>
<td>Text</td>
<td>Digital Resource/Format</td>
<td></td>
</tr>
<tr>
<td><strong>dc:identifier</strong></td>
<td>Heritage Asset Identification/Appellation/ID</td>
<td>Digital Resource/Appellation/ID</td>
<td></td>
</tr>
<tr>
<td><strong>dc:language</strong></td>
<td>Heritage Asset Identification/Record Information/Language</td>
<td>Digital Resource/Language</td>
<td></td>
</tr>
<tr>
<td><strong>dc:publisher</strong></td>
<td>Heritage Asset Identification/Record Information/Source</td>
<td>Digital Resource/Publication statement/publisher</td>
<td></td>
</tr>
<tr>
<td><strong>dc:publisher</strong></td>
<td></td>
<td>Digital Resource/Publication statement/placeOfPublication</td>
<td></td>
</tr>
<tr>
<td><strong>dc:publisher</strong></td>
<td></td>
<td>Digital Resource/Publication statement/date</td>
<td></td>
</tr>
<tr>
<td><strong>dc:relation</strong></td>
<td>Heritage Asset Identification/Relations/Target of the relation</td>
<td>Digital Resource/Relations/Target of the relation</td>
<td></td>
</tr>
<tr>
<td><strong>dc:rights</strong></td>
<td></td>
<td>Digital Resource/Rights</td>
<td></td>
</tr>
<tr>
<td><strong>dc:source</strong></td>
<td>Heritage Asset Identification/Record Information/Source</td>
<td>Digital Resource/Record Information/Source</td>
<td></td>
</tr>
<tr>
<td><strong>dc:subject</strong></td>
<td>Heritage Asset Identification/Characters/Heritage asset type</td>
<td>Digital Resource/Subject</td>
<td></td>
</tr>
<tr>
<td><strong>dc:subject</strong></td>
<td>Heritage Asset Identification/Characters/Craft/Constructionmethod</td>
<td>Digital Resource/Record Information/Keywords</td>
<td></td>
</tr>
<tr>
<td><strong>dc:subject</strong></td>
<td>Heritage Asset Identification/Characters/Craft/Propulsion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dc:subject</td>
<td>Heritage Asset Identification/Record Information/Keywords</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------------------------------------</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>dc:title</td>
<td>Heritage Asset Identification/Appellation/Name</td>
<td>Digital Resource/Appellation/Name</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The preferred name is mapped as a dc:title. If a preferred name is not indicated then the first name is mapped as a dc:title and the following is/are mapped as dcterms:alternative.</td>
<td></td>
</tr>
<tr>
<td>dc:type</td>
<td>Text</td>
<td>Digital Resource/Type</td>
<td></td>
</tr>
<tr>
<td>dcterms:alternative</td>
<td>Heritage Asset Identification/Appellation/Name (not preferred)</td>
<td>Digital Resource/Appellation/Name</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Digital Resource/Created</td>
<td></td>
</tr>
<tr>
<td>dcterms:extent</td>
<td>Heritage Asset Identification/Characters/Dimensions</td>
<td>Digital Resource/Extent</td>
<td></td>
</tr>
<tr>
<td>dcterms:hasPart</td>
<td>Heritage Asset Identification/Relations/Target of the relation</td>
<td>Digital Resource/Relations/Target of the relation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>type of the relation = HasPart</td>
<td></td>
</tr>
<tr>
<td>dcterms:isPartOf</td>
<td>Heritage Asset Identification/Relations/Target of the relation</td>
<td>Digital Resource/Relations/Target of the relation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>type of the relation = isPartOf</td>
<td></td>
</tr>
<tr>
<td>dcterms:isVersionOf</td>
<td>Heritage Asset Identification/Relations/Target of the relation</td>
<td>Digital Resource/Relations/Target of the relation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>type of the relation = isDerivativeOf</td>
<td></td>
</tr>
<tr>
<td>dcterms:medium</td>
<td>Heritage Asset Identification/Characters/Materials</td>
<td>Digital Resource/Medium</td>
<td></td>
</tr>
<tr>
<td>dcterms:provenance</td>
<td>n/a to heritage asset</td>
<td>Digital Resource/Provenance</td>
<td></td>
</tr>
<tr>
<td>Property</td>
<td>Description</td>
<td>Value</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>dcterms:references</td>
<td>Heritage Asset Identification/References (Actors&gt;Nome + Appellation name)</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>dcterms:replaces</td>
<td>Heritage Asset Identification/Relations/Target of the relation (type of the relation = is Successor Of)</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>dcterms:spatial</td>
<td>Heritage Asset Identification/Spatial/Spatial/Location set/Named location</td>
<td>Digital Resource/Spatial/Spatial/Location set/Named location</td>
<td></td>
</tr>
<tr>
<td>dcterms:spatial</td>
<td>Heritage Asset Identification/Spatial/Location set/Address (includes building name, number in road, road name, town or city, postcode/zipcode, locality, admin area, country)</td>
<td>Digital Resource/Spatial/Location set/Address (includes building name, number in road, road name, town or city, postcode/zipcode, locality, admin area, country)</td>
<td></td>
</tr>
<tr>
<td>dcterms:spatial</td>
<td>Heritage Asset Identification/Spatial/Location set/Geopolitical area</td>
<td>Digital Resource/Spatial/Location set/Geopolitical area</td>
<td></td>
</tr>
<tr>
<td>dcterms:spatial</td>
<td>Heritage Asset Identification/Spatial/Location set/Historical name</td>
<td>Digital Resource/Spatial/Location set/Historical name</td>
<td></td>
</tr>
<tr>
<td>dcterms:spatial</td>
<td>Heritage Asset Identification/Spatial/Cartographic reference/Coordinates</td>
<td>Digital Resource/Spatial/Cartographic reference/Coordinates</td>
<td></td>
</tr>
<tr>
<td>dcterms:spatial</td>
<td>Heritage Asset Identification/Characters/Craft/Place of registration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dcterms:spatial</td>
<td>Heritage Asset Identification/Characters/Craft/Nationality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dcterms:spatial</td>
<td>Heritage Asset Identification/Characters/Craft/LastJourneyDetails/Departure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>----------------------------------------------------------------------------</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>dcterms:spatial</td>
<td>Heritage Asset Identification/Characters/Craft/LastJourneyDetails/Destination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dcterms:temporal</td>
<td>Heritage Asset Identification/Characters/Temporal/Period name</td>
<td>Digital Resource/Temporal/Period name</td>
<td></td>
</tr>
<tr>
<td>ens:type</td>
<td>Text</td>
<td>Text, Image, Sound or Video depends on type of resource</td>
<td></td>
</tr>
<tr>
<td>ens:currentLocation</td>
<td>? Heritage Asset Identification/Spatial/Spatial/Location set/Named location</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>ens:isDerivativeOf</td>
<td>Heritage Asset Identification/Relations/Target of the relation (type of the relation = isDerivativeOf)</td>
<td>Digital Resource/Relations/Target of the relation</td>
<td></td>
</tr>
<tr>
<td>ens:isNextInSequence</td>
<td>Heritage Asset Identification/Relations/Target of the relation (type of the relation = isNextInSequence)</td>
<td>Digital Resource/Relations/Target of the relation</td>
<td></td>
</tr>
<tr>
<td>ens:isRelatedTo reference</td>
<td>Heritage Asset Identification/Relations/Target of the relation (type of the relation = isRelatedTo)</td>
<td>Digital Resource/Relations/Target of the relation</td>
<td></td>
</tr>
<tr>
<td>ens:isRepresentationOf</td>
<td>Heritage Asset Identification/Relations/Target of the relation digital resource - id</td>
<td>Digital Resource/Relations/Target of the relation</td>
<td></td>
</tr>
<tr>
<td>ens:isSuccessorOf</td>
<td>Heritage Asset Identification/Relations/Target of the relation (type of the relation = isSuccessorOf)</td>
<td>Digital Resource/Relations/Target of the relation</td>
<td></td>
</tr>
<tr>
<td><strong>edm:WebResource</strong></td>
<td>Digital Resource/Link</td>
<td>Digital Resource/Link</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------------</td>
<td>----------------------</td>
<td></td>
</tr>
<tr>
<td><strong>dc:rights</strong></td>
<td>Digital Resource/Rights/Copyright/Rights holder + Rights dates</td>
<td>Digital Resource/Rights/Copyright/Rights holder + Rights dates</td>
<td></td>
</tr>
<tr>
<td><strong>dc:rights</strong></td>
<td>Digital Resource/Rights/Copyright/Credit line</td>
<td>Digital Resource/Rights/Copyright/Credit line</td>
<td></td>
</tr>
<tr>
<td><strong>ore:Aggregation</strong></td>
<td>Heritage Asset Identification/Record Information /ID or Heritage Asset Identification/Appellation /ID (To be determined)</td>
<td>Digital Resource/Record information/ID</td>
<td></td>
</tr>
<tr>
<td><strong>ore:aggregates</strong></td>
<td>Heritage Asset Identification/Record Information /ID</td>
<td>Digital Resource/Record information/ID</td>
<td></td>
</tr>
<tr>
<td><strong>ens:dataProvider</strong></td>
<td>Heritage Asset Identification/Record Information/Source</td>
<td>Digital Resource/Record Information/Source</td>
<td></td>
</tr>
<tr>
<td><strong>ens:provider</strong></td>
<td>CARARE</td>
<td>CARARE</td>
<td></td>
</tr>
<tr>
<td><strong>ens:hasView</strong></td>
<td>Digital Resource/Link</td>
<td>Digital Resource/Link</td>
<td></td>
</tr>
<tr>
<td><strong>ens:isShownBy</strong></td>
<td>Digital Resource/Link</td>
<td>Digital Resource/Link</td>
<td></td>
</tr>
<tr>
<td><strong>ens:isShownAt</strong></td>
<td>Digital Resource/Link</td>
<td>Digital Resource/IsShownAt</td>
<td></td>
</tr>
<tr>
<td><strong>ens:object</strong></td>
<td>Digital Resource/Object</td>
<td>Digital Resource/Object</td>
<td></td>
</tr>
<tr>
<td><strong>dc:rights</strong></td>
<td>Digital Resource/Rights/Copyright/Credit line</td>
<td>Digital Resource/Rights/Copyright/Credit line</td>
<td></td>
</tr>
</tbody>
</table>
4.1 CARARE Record example
The following example is a record provided by the Scuola Normale Superiore of Pisa.

```xml
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.carare.eu/carareSchema
http://www.carare.eu/carareSchema">
  <car:carare id="">
    <car:collectionInformation>
      <car:title lang="it" preferred="true">La Fortuna Visiva di Pompeii</car:title>
      <car:title lang="en" preferred="false">The visual fortune of Pompeii</car:title>
      <car:keywords lang="it">Pompei</car:keywords>
      <car:keywords lang="en" namespace="http://sws.geonames.org/">Pompeii</car:keywords>
      <car:keywords lang="it">archeologia</car:keywords>
      <car:keywords lang="en" namespace="http://www.getty.edu/research/tools/vocabularies/aat/">archaeology</car:keywords>
      <car:keywords lang="it">monumenti</car:keywords>
      <car:keywords lang="en" namespace="http://www.getty.edu/research/tools/vocabularies/aat/">monuments</car:keywords>
    </car:collectionInformation>
    <car:contacts>
      <car:name>Maria Emilia Masci</car:name>
      <car:role>researcher</car:role>
      <car:organization>Scuola Normale Superiore</car:organization>
      <car:address>piazza dei Cavalieri, 7 - 56026 - Pisa, IT</car:address>
      <car:phone>+39 050 509683</car:phone>
      <car:email>e.masci@sns.it</car:email>
    </car:contacts>
    <car:rights>
      <car:accessRights>
        <car:grantedTo>CCO 1.0 Universal Public Domain Dedication for metadata describing this collection</car:grantedTo>
        <car:conditions>Metadata describing this collection are in the public domain. Attribution of author is required for any copy, derivative work or citation of the present metadata: mention "Scuola Normale Superiore" as author. The database and visual resources related to these metadata are covered by copyright and can be reproduced only if allowed by the rights holder.</car:conditions>
        <car:statement>Author: Scuola Normale Superiore, Pisa, IT</car:statement>
      </car:accessRights>
      <car:reproductionRights>
        <car:conditions>Metadata describing this collection are in the public domain (CCO 1.0 Universal Public Domain Dedication). Attribution of author is required for any copy, derivative work or citation of the present metadata: mention "Scuola Normale Superiore" as author. The database and visual resources related to these metadata are covered by copyright and
```
can be reproduced only if allowed by the rights holder.

Scuola Normale Superiore - La Fortuna Visiva di Pompei

Scuola Normale Superiore, Pisa, IT

2011

2011-01-01 to 1899-12-31

Roma

Età Romana

XVIII and XIX centuries

ante 79 d.C.

1748-1900

Pompeii

80045

Scavi archeologici di Pompei

Campania

Italy

40.75000

14.48333
Documented workflow

23

http://www.geonames.org/maps/google_40.75_14.483.html

Scuola Normale Superiore - La Fortuna Visiva di Pompei

2011-10-18

Scuola Normale Superiore

e.masci@sns.it

Pisa

Author: Scuola Normale Superiore, Pisa, IT

Metadata describing this monument are in the public domain. Attribution of author is required for any copy, derivative work or citation of the present metadata: mention "Scuola Normale Superiore" as author. The database and visual resources related to these metadata are covered by copyright and can be reproduced only if allowed by the rights holder.

Metadata describing this monument are in the public domain (CCO 1.0 Universal Public Domain Dedication). Attribution of author is required for any copy, derivative work or citation of the present metadata: mention "Scuola Normale Superiore" as author. The database and visual resources related to these metadata are covered by copyright and can be reproduced only if allowed by the rights holder.

Thermopolium e caupona di Caprasia e Nymphius

4 SCHELETRI (CASA DEI)
<car:name lang="it" preferred="false">AVELLIUS (HOUSE OF)</car:name>  
<car:name lang="it" preferred="false">AVELLIUS FIRMUS (HOUSE OF)</car:name>  
<car:name lang="it" preferred="false">CANUSIA E NYMPHIUS (CASA DI)</car:name>  
<car:name lang="it" preferred="false">CAPRASIA E NYMPHIUS (CASA E CAUPONA DI)</car:name>  
<car:name lang="it" preferred="false">CAPRASIA E NYMPHIUS (THERMOPOLIUM E CAUPONA DI)</car:name>  
<car:name lang="it" preferred="false">CINQUE SCHELETRI (CASA DEI)</car:name>  
<car:name lang="it" preferred="false">FIRMUS (HOUSE OF)</car:name>  
<car:name lang="it" preferred="false">CAPRASIA E NYMPHIUS (THERMOPOLIUM E CAUPONA DI)</car:name>  
<car:name lang="it" preferred="false">CASA DI CAPRASIA E NYMPHIUS</car:name>  
<car:id>6|10|3,4,18</car:id>  
</car:appellation>  
</car:description>  
</car:characters>  
</car:spatial>  
<car:locationSet>  
<car:namedLocation lang="en" namespace="http://sws.geonames.org/">Pompeii</car:namedLocation>  
<car:address>  
<car:townOrCity authority="http://sws.geonames.org/3170336/about.rdf">Pompeii</car:townOrCity>  
<car:postcodeOrZipcode>80045</car:postcodeOrZipcode>  
</car:address>  
</car:locationSet>  
</car:spatialReferenceSystem>  
<car:quickpoint>  
<car:x>40.75000</car:x>  
<car:y>14.48333</car:y>  
</car:quickpoint>
<car:typeOfRelation>isRepresentationOf</car:typeOfRelation>
<car:targetOfRelation>6029</car:targetOfRelation>
</car:relations>
<car:rights>
    <car:copyright>
        <car:rightsHolder>Firenze, Biblioteca Nazionale Centrale - su concessione del Ministero per i beni e le attività culturali</car:rightsHolder>
        <car:creditLine>© Firenze, Biblioteca Nazionale Centrale - su concessione del Ministero per i beni e le attività culturali</car:creditLine>
    </car:copyright>
    <car:accessRights>
        <car:statement>© Firenze, Biblioteca Nazionale Centrale - su concessione del Ministero per i beni e le attività culturali - all rights reserved</car:statement>
    </car:accessRights>
</car:rights>
</car:carare>
</car:carareWrap>
4.2 EDM record example

The following example is a transformation of the record shown in 5.1 above to EDM format.

```xml
<edm xmlns=""
    xmlns:ens="http://pompeii.sns.it/prado"><ens:ProvidedCHO rdf:about="6029">
    <dc:date xml:lang="en">before 79 a.C.</dc:date>
    <dc:date xml:lang="it">ante 79 d.C.</dc:date>
    <dc:description xml:lang="en">regio: 6, insula: 10, civico: 3,4,18.</dc:description>
</ens:ProvidedCHO>
<dc:date xml:lang="en">before 79 a.C.</dc:date>
<dc:date xml:lang="it">ante 79 d.C.</dc:date>
<dc:description xml:lang="en">regio: 6, insula: 10, civico: 3,4,18.</dc:description>
</edm>
```

The following example is a transformation of the record shown in 5.1 above to EDM format.

```
<edm xmlns=""
    xmlns:ens="http://pompeii.sns.it/prado"><ens:ProvidedCHO rdf:about="6029">
    <dc:date xml:lang="en">before 79 a.C.</dc:date>
    <dc:date xml:lang="it">ante 79 d.C.</dc:date>
    <dc:description xml:lang="en">regio: 6, insula: 10, civico: 3,4,18.</dc:description>
</ens:ProvidedCHO>
<dc:date xml:lang="en">before 79 a.C.</dc:date>
<dc:date xml:lang="it">ante 79 d.C.</dc:date>
<dc:description xml:lang="en">regio: 6, insula: 10, civico: 3,4,18.</dc:description>
</edm>
```

The following example is a transformation of the record shown in 5.1 above to EDM format.

```
<edm xmlns=""
    xmlns:ens="http://pompeii.sns.it/prado"><ens:ProvidedCHO rdf:about="6029">
    <dc:date xml:lang="en">before 79 a.C.</dc:date>
    <dc:date xml:lang="it">ante 79 d.C.</dc:date>
    <dc:description xml:lang="en">regio: 6, insula: 10, civico: 3,4,18.</dc:description>
</ens:ProvidedCHO>
<dc:date xml:lang="en">before 79 a.C.</dc:date>
<dc:date xml:lang="it">ante 79 d.C.</dc:date>
<dc:description xml:lang="en">regio: 6, insula: 10, civico: 3,4,18.</dc:description>
</edm>
```

The following example is a transformation of the record shown in 5.1 above to EDM format.

```
<edm xmlns=""
    xmlns:ens="http://pompeii.sns.it/prado"><ens:ProvidedCHO rdf:about="6029">
    <dc:date xml:lang="en">before 79 a.C.</dc:date>
    <dc:date xml:lang="it">ante 79 d.C.</dc:date>
    <dc:description xml:lang="en">regio: 6, insula: 10, civico: 3,4,18.</dc:description>
</ens:ProvidedCHO>
<dc:date xml:lang="en">before 79 a.C.</dc:date>
<dc:date xml:lang="it">ante 79 d.C.</dc:date>
<dc:description xml:lang="en">regio: 6, insula: 10, civico: 3,4,18.</dc:description>
</edm>
```
Roman

Età Romana

Regio VI, 8 e Case VI, 10 che si affacciano sulla Via di Mercurio - pianta | incisione | acquaforte

William Gell, Pompeiana: The Topography, Edifices and Ornaments of Pompeii. The Result of Excavations since 1819, 1832, Tav. LX
D 3.3 Documented workflow
Annex II. Mapping and Implementation Guidelines for the CARARE metadata schema

CARARE provides Europeana with an aggregation service to deliver access to digital objects and information resources relating to the archaeological and architectural heritage. This includes a diverse range of materials relating to unique monuments, buildings, landscapes, heritage sites and artefacts provided by national heritage agencies, research organisations and archaeological museums across Europe. The aim is to provide integrated access to the archaeological and architectural heritage in a common online environment.

Each CARARE content provider has created its digital resources and metadata following its own organizational principles, descriptive standards and management procedures. Although the international conventions on conservation of the archaeological and architectural heritage provide common ground between the partner organisations and their information resources, a very diverse range of cultural assets and media formats are being described and there is no universal metadata standard or ontology which has been applied by all of the organisations involved in the network.

CARARE has established a metadata schema to use as a domain specific harvesting protocol in order to mediate between the native metadata collected by its partners and Europeana. This schema defines a set of standard elements which are based on existing standards from the archaeology and architecture domain. The CARARE metadata schema was released in autumn 2010 and updated in spring 2011 following testing by content partners. It is a harvesting schema which is based on MIDAS Heritage, LIDO and the CIDOC CRM.

To use the CARARE schema it is necessary for partners to map elements in their native metadata to the CARARE metadata format. In addition to the mapping it may be necessary for some normalisation or metadata enrichment processes to be carried out to improve machine readability. Once the mapping and normalisation has been carried out, native metadata can be transformed into CARARE schema format and ingested to the CARARE repository ready for supply to Europeana.

The CARARE aggregator has undertaken to carry out the work needed to transform CARARE records into Europeana’s preferred metadata format; currently this is an implementation of the Europeana Data Model (EDM version 5.2) designed for the Danube release of Europeana.

The mapping and implementation guidelines for the CARARE metadata schema are available in the file: Implementation-guidelines-CARARE-metadata-schema.pdf.