A New Architecture and Approach to Asset Representation for Europeana Aggregation: The CARARE Way

Christos Papatheodorou[1,2], Costis Dallas[1,3], Christian Ertmann-Christiansen[4], Kate Fernie[5], Dimitris Gavrilis[1], Maria Emilia Masci[6], Panos Constantopoulos[1,7], and Stavros Angelis[1]

[1] Digital Curation Unit – IMIS, Athena Research Centre, Athens, Greece
[2] Department of Archives and Library Sciences, Ionian University, Corfu, Greece
[3] Department of Communication, Media and Culture, Panteion University, Athens, Greece
[4] Heritage Agency of Denmark, Copenhagen, Denmark
[5] MDR Partners, UK
[6] Laboratorio LARTTE, Scuola Normale Superiore di Pisa, Italy
[7] Department of Informatics, Athens University of Economics and Business, Athens, Greece
{c.papatheodorou,c.dallas,d.gavrilis,p.constantopoulos, s.angelis}@dcu.gr, chrert@kulturarv.dk, kfernie@tiscali.co.uk, e.masci@sns.it

Abstract. This paper presents a new metadata aggregation approach based on a mediating repository that intends to ensure the integrity, authenticity and semantic enrichment of metadata provided to Europeana by heterogeneous collections. Primary metadata are mapped to CARARE schema, a schema suitable for describing archaeological and architectural heritage assets, digital resources, collections, as well as events associated with them. The paper specifies the proposed schema and discusses the overall architecture of the proposed approach.

Keywords: Cultural heritage metadata, Composite content, Metadata integration, Infrastructures for sharing content.

1 Introduction

Archaeological and architectural monuments and sites constitute an important part of tangible European cultural heritage. Aligned with the initiative to establish the Europeana digital library as a single point of access to Europe's cultural heritage, the CARARE Best Practice Network was established in 2009 in order to increase the quantity and quality of digital content for the archaeological and architectural heritage that is available to users of Europeana, while addressing some significant issues specific to archaeological and architectural heritage. An overview of the CARARE project, its workplan and the challenges it purports to meet is presented in [1].

The objective of CARARE is to integrate in Europeana an estimated 2 million digital resources, corresponding to approximately 1 million unique monuments, buildings, landscapes, heritage sites and artefacts. Content belongs to a broad and diverse set of heritage organizations across Europe, each adhering to different organizational

E. García-Barriocanal, A. Öztürk, and M.C. Okur (Eds.): MTSR 2011, CCIS 240, pp. 412–423, 2011.

principles, descriptive standards and management procedures with regard to their data. The cultural assets to be made available are very diverse, from prehistoric and Iron age archaeological survey results to complex Mediterranean archaeological sites, and historic buildings. The digital resources representing such assets are also heterogeneous, ranging from paintings and prints to photographs, archaeological and architectural plans, sections and drawings, and, increasingly, digital 3D models and renderings thereof [2]. Heritage assets are as a rule associated with geographic information, both in the form of geographic coordinates according to some grid standard, and in the form of named geographic entities such as historical locations and areas. In addition, some heritage assets – for instance, an archaeological site such as Pompeii – are characterized by a nested mereological structure, being composed of buildings, each of which is also composed by particular architectural elements; resources may, indeed, represent architectural and archaeological assets at quite different levels of complexity. In general digital resources are as a rule associated with a heritage asset but in some cases the monuments and buildings captured in historical images are no longer known.

Earlier initiatives to provide metadata and links to digital resources for aggregation in Europeana are based on a mapping architecture, by virtue of which original metadata records are translated – following a set of atomic rules for individual resources and data elements – into a common output schema such as the Europeana Semantic Elements (ESE) [3]. CARARE represents a significant departure from this architecture. It introduces the notion of an information broker – an intermediate repository acting as a mediator– intended to ensure the integrity, authenticity and content enrichment of metadata provided to Europeana by heterogeneous collections. This paper presents the rationale, the background, as well as the specific requirements in the field of schema definition, overall architecture and technical approach, as well as content enrichment, adopted by CARARE to meet the challenges presented above.

2 Background

Lightweight Information Describing Objects (LIDO) [4] is an output of the work of CIDOC Working Group Data Harvesting and Interchange and the result of a joint effort by the CDWA Lite - museumdat Working Group, the SPECTRUM community, the Documentation Committee of the German Museums Association and the ATHENA project. LIDO is intended to harvest information for museum objects, and as such it aims to aggregate information from several standards and schemes that have been developed in Europe and the rest of the world. These include the following schemes and standards: CDWA Lite, CIDOC CRM, museumdat and SPECTRUM. Its perspective is event-oriented, a view originating from the contribution of CIDOC CRM [5]. The schema provides seven areas in each record for an object. Each metadata record includes (a) a Metadata Record ID, (b) a Category, (c) the descriptive metadata and (d) administrative metadata of the record.

On the other hand, MIDAS Heritage [6] is a standard, which indicates the kind of information that should be recorded to support effective sharing and long-term preservation of the knowledge of the historic environment. MIDAS has a three-level structure: (a) Themes, which are the highest level information blocks of interest for

the historic environment community, (b) Information Groups, which set the specific standard for what should be included in an entry covering a particular subject and (c) Units of Information, which are the basic items that make up an entry.

The main conclusions of a comparative study between LIDO and MIDAS schemas are that MIDAS is a more general standard covering several areas of Cultural Heritage Assets Management, while LIDO focuses mainly on the description of museum objects, providing a semantically rich carrier for information exchange. MIDAS organizes the Cultural Heritage Assets to the classes Monuments, Areas and Artefact-Ecofact and provides elements for the rich description of their geospatial features, without focusing on the digital representations. It provides a significant set of elements for cartography. The added value feature of MIDAS is that it provides analytical information on the interventions, works and studies referring to a Cultural Heritage Asset. Moreover the Theme Information Sources allows for the rich documentation of the objects and the interventions on them. LIDO provides elements for the detailed description of museum objects, as well as their digital representations. Since it focuses on museum objects, it does not cover significant aspects of the cultural heritage assets documentation, such as their geographical identification and the history of the interventions and restorations. In general, the two schemas have different orientations and serve different purposes and functions. MIDAS is more general and attempts to cover a wide range cultural heritage asset types as well as the activities performed on them, while LIDO provides a descriptive expressiveness for museum objects. Although both schemas share many common entities, these are interpreted differently according to the orientation and scope of each schema.

A deeper comparison between the two schemas is needed to reveal the elements needed to constitute the CARARE schema. This comparison should be based on CIDOC CRM ontology, due to its power to express clearly the semantics of the cultural heritage metadata schemas [2]. In particular the study and comparison of the schemas with the CIDOC CRM, is based on the POLIS DTD [7], a DTD derived directly by the CIDOC CRM. The POLIS DTD was produced as part of an EU funded Greek national research project to develop an interoperability framework for the cultural heritage. It provides a language to describe immovable monuments, a domain that covers adequately the domain of CARARE project. It focuses on both the objects and their digital representations, but it is a descriptive schema and therefore it is not so analytical in handling information concerning the management of the cultural heritage assets, such as the description of particular investigative, consultation, research, etc. activities. However due to its event-based character, it is powerful and flexible enough to handle the information for these activities as events.

Given that the two metadata schemas include a lot of elements and the comparison of their semantics requires a lot of effort, the existence of a DTD, which directly implements particular CIDOC CRM expressions covering almost completely the CARARE domain, facilitates significantly the process of the comparison that aims to the development of an appropriate schema. The added-value features of the approach we followed are: (a) homogeneity: the metadata schemas are compared and discussed under the prism of another metadata schema, (b) facility: there is no need to develop all the CIDOC CRM expressions and to correlate them with LIDO and MIDAS schemas. The mapping emerged the following conclusions: (i) The proposed structure of the CARARE schema should be an application profile based on MIDAS, which

covers several activities for the Cultural Heritage Assets Management and ensures metadata provenance. Moreover MIDAS is almost completely aligned (mapped) to CIDOC CRM paths, while the elements that are not covered by the schema are considered of minor importance (ii) LIDO provides a powerful framework for the detailed description of the digital resources that correspond to heritage assets and the related activities and interventions on them.

3 CARARE Schema

The CARARE schema is designed to support the delivery of metadata to the CARARE aggregator and to Europeana. It is an application profile based on MIDAS Heritage, LIDO and the POLIS DTD and the rationale for this decision comes from the nature of the problem it is designed to solve.

CARARE is aggregating content for Europeana from a wide variety of organizations and digital sources with varied and particular histories, which yet draw on common practices that have been developed over the years and both have informed and have been informed by standards, in particular the CIDOC CRM and the CIDOC core data standard for archaeological sites and monuments. These sources include, e.g., administrative/scientific national registries of sites and monuments, archaeological museum collections, collections of 3D models describing any of these types of objects, as well as digital historical document collections such as the Visual Fortune of Pompeii archive described below. At the same time, all of these sources must be aggregated and delivered into the common format, which Europeana now uses to describe its content, the Europeana Data Model (EDM) [8], a target which is developing and changing over time.

Implementing the CARARE schema has followed an iterative approach with an initial adaptation to account for the particular focus of Europeana on digital resources, and the implementation of the schema followed an iterative approach grounded in specific real cases as detailed in the next section.

The CARARE schema's focus is on the detailed description of monuments, digital resources and related events in which the monument has been involved. The Schema follows the structure of MIDAS with the inclusion of elements from the LIDO Resource Set and ESE (Europeana Semantic Elements) to cover the information needed for the digital resources being made accessible to the CARARE and Europeana service environments. Conceptually the top-level themes in the CARARE schema are:

- Heritage asset Identification – the descriptive information and metadata about the monument, historic building, archaeological landscape area, shipwreck, artifact, ecofact etc.
- Digital resource – these are digital representations and sources of information (images, texts, videos, audio, 3D models) about the heritage asset being provided to the service environment (e.g. Europeana). They are often digital representations of monuments or of parts of monuments.
- Activity – these are events that the heritage asset has taken part in, such as Creation, Field investigation, Research and analysis Historical events, etc.
- Collection – this is a collection level description of the data being provided to the service environment.

An XSD has been developed to support the harvesting of metadata from the CARARE content providers to the CARARE aggregator for supply to Europeana. In the CARARE XSD each of these themes are global wrappers which are contained within a CARARE wrapper. An object in the CARARE schema consists of the Heritage Asset Identification (HA) wrapped together with the related Digital Resources (DR), Activities (A) and Collection information (C) (Figure 1). The cardinality of themes and elements has been specified to enable the harvesting of the real data actually present in CARARE content providers' datasets and to meet Europeana's requirements. Some CARARE content providers have collections of historical images where the heritage asset which was photographed is unknown. These collections are interesting for Europeana and for the study of archaeology and archictecture in a region. Since Digital Resources are required for Europeana, the XSD specifies that it is mandatory to include at least one Digital Resource in a CARARE record, while it is highly recommended that each CARARE record should include a single Heritage Asset.

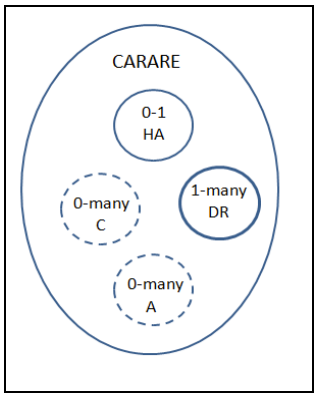

Fig. 1. Representation of a CARARE object containing zero or one heritage asset and its related one to many digital resources, zero to many collection descriptions and zero to many activities

For each of the 4 top-level themes the CARARE schema defines a rich set of descriptive information and also administrative metadata. A series of global information types are used globally across the schema to define elements including:

- Appellation – the name (or title) and identification number
- Description – a textual description of the features of the asset
- Spatial – place name data, geo-spatial coordinates and other spatial data.
- Temporal – dating or other information relating to the time span.
- Actors – data about people, organisations or groups and their roles.
- Address – postal address
- Rights – associated with the object or metadata
- Relations – between the top-level themes i.e. heritage asset, activity, digital resource and collection
- Record information – metadata about the record.

In addition to these global elements, the Heritage asset identification element wraps the following additional elements:

Designations - information about any designations for a monument or building which provide it with protection in law. Its sub-elements are Protection type, Grade (the level of protection), and related dates.

Conditions - information about the condition of a monument or building. Its sub-elements comprise a detailed assessment of the condition of a Heritage Asset, any treatment required and an estimation of the percentage of the monument affected, as well as the date when the assessment of condition was made.

Characters - it wraps the main descriptive metadata of the asset and comprises:

- Heritage asset type: classification of the monument, building, landscape feature, artefact or ecofact primarily with respect to its function or use, e.g. house.
- Materials: the basic materials of which a monument is composed, e.g. brick, stone, tile.
- Inscriptions: text inscribed on a monument or building, if any. The element may be repeated using the xml:lang attribute if the element value is available in alternate languages. A preferred/alternate attribute may be used to indicate which value is preferred. The type of inscription may be indicated using an attribute.
- Dimensions: Holds the measurement type (e.g. height, length, width, depth), shape of the heritage asset (e.g. oval) as well as the measurement units (e.g. metres, centimetres), scale and of course the corresponding values.
- Repository location: It identifies the institution with custody of the artifact and possibly the current location.

In addition to the global elements the information about a digital resource includes:

- Format – the file format of the resource, e.g. MIME type.
- Subject – the subject or topic of the resource.
- Publication statement **–** the name of the publisher, place and date of publication
- Type – The nature or genre of the resource.
- Link– the URL of the resource. A reference to the digital object on the content provider's web site in the best available resolution/quality (i.e. a link to the resource as a text, image, sound, or video file, not to the webpage that contains it). The data given here will allow the automatic generation of a thumbnail by Europeana for its functionality.
- Object – A URL to a thumbnail.
- IsShownAt – A URL to the digital object on the content provider's website in its full information context (i.e. a link to the webpage that contains the digital object and contextual information).
- Resource metadata location – pointer to other information about the resource making the resource available.
- Rights – the rights associated with the digital object itself (copyright, access rights, reproduction rights).

An outline of the CARARE schema together with its implementation in an XSD was published on the project website, and was implemented for testing by project partners in the CARARE metadata mapping and ingestion tool. Following testing some

adaptations were made to the implementation of the schema in the XSD, including for example definition of the mandatory and repeating elements, definition of the XML type of the elements (text, date, integer etc).

4 Testing and Improvements

The definition of the CARARE schema is the result of an interactive process among metadata experts and content providers. In fact, a collaborative testing phase based on the mappings from the original databases provided by the CARARE Partners and the CARARE Application Profile, led to its progressive modification and improvement. Such methodology ensures a strong applicability of the CARARE schema to the domain chosen for the aggregator, mainly centered on archaeological and architectural sites heritage.

The main mapping test was carried out on contents provided by the Scuola Normale Superiore of Pisa (hereafter SNS), which participates in the project both as content provider and as metadata expert, because it will provide metadata describing 2D images and 3D models related to archaeological monuments, thus offering a complete example. Metadata provided by SNS will describe resources from the following two inter-related collections:

- The Visual Fortune of Pompeii: a digital archive of ancient visual documents representing the monuments of Pompeii [9];
- The 3D model of Pompeii Civil Forum: a multi-resolution survey documenting the present status of the buildings in the Forum of Pompeii, carried out in 2009 by the Politechnique of Milan [10].

In the first case, contents of the Visual Fortune of Pompeii database have been directly mapped to the CARARE schema. In the second case, 3D models are shown through 3D PDF documents and further enriched with explicative texts, then metadata describing both the 3D PDF documents and the survey activity are created ad hoc.

The Visual Fortune of Pompeii database is articulated as follows:

- Iconographic sources: entity describing 2D images (ancient prints, drawings and photographs) representing the monuments of Pompeii;
- Bibliographic sources: entity describing digitized ancient books from which are taken the Iconographic sources;
- Archival sources: entity describing digitized manuscripts and inedited documents from which are taken the Iconographic sources;
- Places: authority file describing each of the monuments of Pompeii;
- Persons: authority file describing physical and moral persons related to the various types of sources with different roles (e.g. invention, creation, publication);
- Subjects: thesaurus of subjects represented in the Iconographic sources.

The above listed entities and authority files are related to each other with various bi-directional relations. In particular, Iconographic sources (2D images) and Places (monuments) are linked with a many-to-many relation, as one monument can be represented in many 2D images, and one 2D image can represent many monuments. In addition, as previously pointed out, metadata will also be created to describe: 3D PDF

documents including 3D models and related texts on a given monument; and activities of survey and creation of 3D models representing the monuments.

A modeling issue encountered during the initial testing was how to shape a CARARE record that is which elements to include, how to manage elements cardinality and their reciprocal relations. The initial CARARE schema outline allowed many-to-many relation among monuments (Heritage Assets) and 2D/3D digital resources. This allowed different possible approaches: (a) to include within a CARARE record one monument and the related digital resources; (b) to include only one digital resource and the related monuments; (c) to wrap within one record all of the monuments and all of the related digital resources.

Approaches (a) and (b) cause repetition of information: in the first case information on digital resources can be duplicated in many records, while in the second case there is a duplication of information on monuments. Approach (c) solves the duplication problems, but the risk is to include an entire collection within a single record. After an accurate evaluation, as the CARARE repository is focused on heritage as-sets and monuments, approach (a) was chosen. This choice resulted in the refining of the CARARE schema and the architecture of the system as follows: duplications will be managed within the repository at the DIP level, when packaging contents for Europeana. This also resulted in the decision to include mandatory unique identifiers within each of the CARARE sub-elements, and the rule to represent exclusively direct relations, assuming that inverse relations are always inferred, in order to better manage bi-directional relations. The resulting high-level mapping is resumed in table 1.

Table 1. High-level mapping of CARARE to SNS metadata

Cardinality	SNS contents	CARARE elements
1	4. Places (monuments)	heritageAsset
0-to-N	2. Bibliographic sources 3.Archival sources	heritageAsset/ references
0-to-N	1. Iconographic sources 7. 3D PDF documents	digitalResource
0-to-N	8. Survey and 3D model creation	Activity

Physical and moral persons (n. 5) are mapped within various elements in the sub-element <actors>, using different <roles>. The testing also resulted in the addition of attributes to elements in the XSD. For instance, the availability of SNS contents and of other contents in more than one language led to the addition of the "xml:lang" attribute to be added to all text elements to allow alternate language labels to be identified; the main language of the metadata is specified within the record information sub-element <language>. In this way title, description, keywords and other elements are multi-lingual. Another example is the choice to allow for repeatable <title> sub-elements, with an attribute to indicate which of the titles is "preferred". This is due to the fact that different titles and translated titles are often used for cultural digital resources, and monuments, as it came out from the mapping tests.

In conclusion, the testing phase of the CARARE schema was carried out on real examples from content providers, brought to add modifications and improvements to

the schema and to draw up practical guidelines for its usage, leading to a stable and consistent version of the XSD, which effectively responds both to the objectives of the project and to the real situations of content providers.

5 CARARE Architecture and the MoRe System

Typical Europeana aggregation projects involve the use of a metadata mapping system and an OAI-PMH export system to map native metadata schemas to Europeana Semantic Elements (ESE) schema and recently the Europeana Data Model (EDM) and then expose it through OAI-PMH for harvesting. Despite the fact that this approach is tested and is robust, due to its simple architecture it lacks the added value services that can be provided in this kind of projects. The CARARE architecture (Figure 2) introduces a central repository where all metadata are stored and enriched before being mapped to EDM and provided to Europeana. The concentration of all metadata objects into one system creates added value by allowing the content providers to enrich their objects using semantic relations, measure the quality of their metadata, etc. The use of a central system also facilitates the preservation of the providers' metadata since it holds not only the transformed metadata but also the native.

MoRe, the CARARE repository expands the MOPSEUS [11] functionalities, which is a fedora-commons based repository, fully compliant to OAIS. The repository allows the ingestion of metadata not only through traditional PMH based methods but also using submission packages that contain complex datastreams and information in order to preserve as much information as possible and also to allow for the implementation of added value services. All information that exists in the repository takes the form of three kinds of packages: submission, archival and dissemination as defined by the OAIS model. The various services use this information and create their own packages, which are then ingested into the repository by creating new versions.

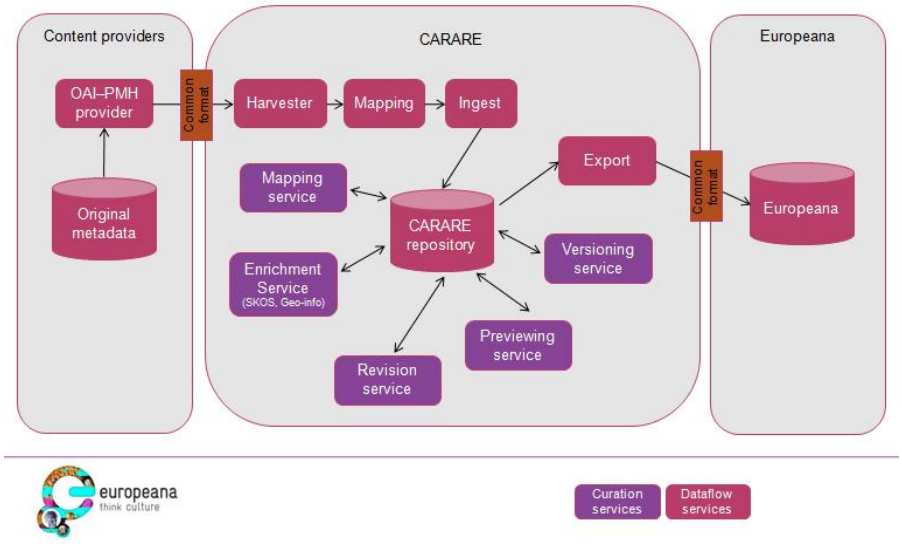

Fig. 2. The CARARE system architecture

MoRe provides a series of services for CARARE: geo-information normalization, semantic enrichment, metadata completeness monitoring, etc. The main functionality of MoRe: semantic enrichment, is a process where objects are enriched using semantic relations derived from EDM. The user is provided with the means to quickly locate relevant objects and add semantic relations between them. During the enrichment process, the enriched metadata datastreams are created in order to preserve the whole information that exists inside MoRe. Other services provide the users with the means to quickly locate geographical and temporal information that exists in their objects, and see how this information will be parsed by Europeana services, normalize it if necessary etc. The constant monitoring and reporting of metadata completeness ensures the quality of metadata delivered to Europeana by performing checks that go beyond the traditional mandatory field checks.

6 Discussion

CARARE schema builds on a long history of schema development amongst the archaeology/architecture heritage community. It is influenced by early work on the CIDOC core data standard for archaeological sites, by the CIDOC CRM, by work on LIDO and the Europeana Data model. It has information resources about monuments and buildings, and their digital representations at its centre.

Place, or the location of a monument or building, is one of the most frequent starting points for enquiries about the archaeological and architectural heritage and is important for both research and preservation of the historic environment. Monument inventories normally include the details of both named locations and geospatial coordinates to enable sites to be located in the real world. The availability of geospatial data is allowing heritage datasets to be included in geo-portals used in environmental planning and monuments and buildings to be viewed alongside the details of other protected resources. The CARARE metadata schema provides a rich set of spatial data elements covering named places (such as Stonehenge or the temple of Dion) and geospatial coordinates (including points, bounding boxes, height and other aspects). The widely avail-able place name gazetteers typically stop at the level of the names of cities, villages, administrative areas, and the more famous buildings and monuments. This means that CARARE has the potential to develop linked data for place-names and to contribute to the development of geo-parsers used by Europeana to enrich records which include place names in free text but lack spatial coordinates.

Realizing this potential involves addressing an issue posed by the native systems used by the CARARE content providers which is specific to geographic data that is the use of national map reference systems. Integration of the data into an international service such as Europeana requires the conversion of geographic data from the national coordinate systems (such as the OS national grid system in use in the UK) to a world geodetic system (such as WGS84). Converting the coordinates in this way enables them to be plotted on world maps and provides a consistent base for information retrieval. CARARE will offer a service to its content providers to convert their coordinates to WGS84 after data has been ingested to the CARARE repository and before it is provided to Europeana.

The mapping of CARARE schema to Europeana Data Model (EDM) [8] was a complex process since the former is an XML Schema, while the latter is a conceptual

model expressed by OWL and RDF Schema. Europeana has recently released an XML Schema for the EDM, for data providers to use when uploading metadata based on it. Our conceptual mapping approach adopted the following rules:

- The CARARE elements are mapped either directly to EDM Classes and ESE elements, or to particular EDM paths. An EDM path is defined as a sequence of Class /subclass /ESE Element -> EDM Property -> Class /subclass /ESE Element.
- We consider that a monument is an instance of the EDM class Physical Thing. A monument is identified by a set of particular characteristics that refer to its nature, location, related events, etc. Thus these characteristics are grouped under the class Physical Thing.
- The information about a Heritage Asset or a digital resource is an instance of the class Europeana Object. This information could be either textual metadata (such as title, etc.), thumbnails and other digital objects representing the monument.

The XML Schema for the EDM which was recently released by Europeana provides the framework for its implementation in the Europeana interface in autumn 2011. The exclusion of the classes 'Physical Thing' and 'Activity' from this implementation has implications for CARARE. Our conceptual mapping of Heritage Asset Identification information and of Digital resources to the class Europeana Object (or Provided Cultural Heritage Object) provides the framework for data supply to Europeana. While the maintenance of the full metatada (including data about Activities and other elements which cannot be mapped to EDM) on the CARARE MoRe repository will facilitate future developments.

7 Conclusions

In this paper we presented the innovative approach introduced by the CARARE Best Practice Network to the aggregation of metadata for architectural and archaeological assets, and related digital resources, in Europeana. The CARARE approach is based on the introduction of MoRe (Monument Repository), an intermediate, OAIS-compliant, metadata aggregator acting as a mediator between heterogeneous primary collections and the Europeana digital library. Primary resources are mapped to CARARE schema, a common metadata schema suitable for the representation of metadata describing heritage assets, digital resources, collections, and events associated with them, thus encompassing adequately the scope of information typically found in relevant primary collections. The MoRe repository manages ingested metadata (including those related to 3D digital resources) and related geographic information in an OAIS-compliant way, providing for the semantic enrichment, long-term preservation and curation [12] of metadata, and for their flexible delivery to emerging new standards within the Europeana ecosystem, such as EDM.

References

1. Hansen, H.J., Fernie, K.: CARARE: Connecting Archaeology and Architecture to Europeana. In: Ioannides, M., Fellner, D., Georgopoulos, A., Hadjimitsis, D.G. (eds.) EuroMed 2010. LNCS, vol. 6436, pp. 450–462. Springer, Heidelberg (2010), http://www.springerlink.com/content/978-3-642-16872-7

2. Cruz, I.F., Xiao, H.: The role of ontologies in data integration. Engineering Intelligent Systems 13(4), 245–252 (2005)
3. Europeana v.01: Europeana Semantic Elements Specifications, ver. 3.2.2 (2010), http://www.version1.europeana.eu/c/document_library/get_file?uuid=c56f82a4-8191-42fa-9379-4d5ff8c4ff75&groupId=10602
4. Data Harvesting and Interchange Working Group: LIDO - Lightweight Information Describing Objects (2009), http://www.cidoc.icom/museum/WG_Data_Harvesting(en)(E1).xml
5. ICOM/CIDOC Documentation Standards Group, CIDOC CRM Special Interest Group: Definition of the CIDOC Conceptual Reference Model (2011), http://cidoc.ics.forth.gr/definition_cidoc.html
6. English Heritage: MIDAS Heritage: the UK historic environment data standard. English Heritage, 2nd edn. (2007), http://www.english-heritage.org.uk/
7. Constantopoulos, P., Bekiari, C., Doerr, M.: POLIS DTD (2005) (in Greek), http://www.ics.forth.gr/CULTUREstandards/paradotea/paradotea_final/K7_odigos_akiniton_perigrafi_2006.pdf
8. Europeana v.01: Definition of the Europeana Data Model elements ver. 5.2.2 (2011), http://version1.europeana.eu/web/europeana-project/technicaldocuments/
9. La Fortuna visiva di Pompei, http://pompei.sns.it
10. Guidi, G., Remondino, F., Russo, M., Rizzi, A., Voltolini, F., Menna, F., Fassi, F., Ercoli, S., Masci, M.E., Benedetti, B.: A multi-resolution methodology for archaeological survey: the Pompeii Forum. In: Proc. VSMM 2008 – Conference on Virtual Systems and Multi-Media dedicated to Digital Heritage. Project papers, pp. 51–59 (2008)
11. Gavrilis, D., Papatheodorou, C., Constantopoulos, P., Angelis, S.: Mopseus – A Digital Library Management System Focused on Preservation. In: Lalmas, M., Jose, J., Rauber, A., Sebastiani, F., Frommholz, I. (eds.) ECDL 2010. LNCS, vol. 6273, pp. 445–448. Springer, Heidelberg (2010)
12. Angelis, S., Benardou, A., Constantopoulos, P., Dallas, C., Gavrilis, D.: A curation-aware repository supporting information-intensive scholarly research. In: 6th Intl. Digital Curation Conference: Participation & Practice: Growing the Curation Community Through the Data Decade (2010)