

3D-ICONS METADATA SCHEMA FOR 3D OBJECTS^{*}

Andrea D'Andrea, Università degli Studi di Napoli "L'Orientale" (UNO)

Kate Fernie, MDR Partners, London (MDR)

Introduction

The paper synthetically reports about the solutions adopted in the first year of the 3D-ICONS project as far as the metadata are concerned and other quality recommendations have been implemented. One of 3D-ICONS main aims was to prepare an extension to the CARARE schema to support provenance and paradata required for the quality assurance of the 3D models. The project also intended to contribute to the maintenance of the CARARE schema by monitoring developments in the core module of the Europeana metadata schema (EDM).

In Section 2 the paper focuses on the description of the starting point of the project in terms of definition of metadata for 3D objects; it introduces the CARARE metadata schema and the EDM schema highlighting the state-of-art of the two schemas showing some recent case-studies based on the integration of the two metadata schemas (CARARE-EDM). As far as the provenance is concerned, the paper describes the CRM_{dig} schema recently adopted and customized by the 3D-COFORM project. Finally it addresses the *paradata principles* and how the input of paradata in the metadata schema are relevant for the new strategy of Europeana. Sections 3 and 4 report about the definition of the CARARE 2.0 schema based on the objectives of the 3D-ICONS and the results achieved in the mapping of CARARE on the EDM schema. In particular the paper outlines how to model provenance and paradata concepts implementing an event approach and correlating different activities that the object has taken part in.

* The paper is largely based on the deliverable D6.1 on Metadata and Thesauri submitted at the end of the first year of the 3D-ICONS Project: <http://3dicons-project.eu/eng/Media/Files/D6.1-Report-on-Metadata-Thesauri>. The deliverable has been prepared by UNO with the support of other partners: the authors are Andrea D'Andrea (UNO) and Kate Fernie (MDR). For the general objectives of the 3D-ICONS project (D'Andrea 2012, 87-109).

*The state of art: EDM and CARARE schemas*EDM Schema

The EDM¹ is a qualitative change in the way Europeana deals with the metadata gathered from data providers and aggregators. It is aimed at solving some of the issues observed with the ESE by providing extra expressivity and flexibility. In particular, EDM makes a distinction between the intellectual and technical creation that is submitted by a provider (a bundle of resources about a digital object created by the provider), the object this structure is about, and the digital representations of this object, which can be accessed over the web. Also, EDM adheres to the modelling principles that underpin the approach of the Web of Data. In that approach, there is no such thing as a fixed schema that dictates just one way to represent data. A common model like EDM can be seen instead as an anchor to which various finer-grained models can be attached, making them at least partly interoperable at the semantic level, while the data retain their original expressivity and richness. It does not require changes in the local approaches, although any changes in the local practice that increase the cross-domain usefulness of the data is encouraged, such as the use of publicly accessible vocabularies (for persons, places, subjects etc.). In this sense, EDM is an attempt to transcend the respective information perspectives of the various communities constituting Europeana, such as museums, archives, audio-visual collections and libraries. EDM is not built on any particular community standard but rather adopts an open, cross-domain framework that can accommodate particular community standards.

As EDM² is intended to be an integration medium for collecting, connecting and enriching the descriptions provided by Europeana content providers, it can include any element (i.e., class or property) found in a content provider's description. Giving an account of all these elements is clearly an impossible task, since they form an open set, i.e. a set that can be extended as new providers join the Europeana information space.

¹ “*Europeana Data Model Primer*”, (26/10/2011) Europeana v1.0,

<http://pro.europeana.eu/documents/900548/770bdb58-c60e-4beb-a687-874639312ba5>

² “*Definition of the Europeana Data Model elements*” Version 5.2.3, (24/02/2012) Europeana v 1.0, <http://pro.europeana.eu/documents/900548/bb6b51df-ad11-4a78-8d8a-44cc41810f22>

EDM has three core classes of resources that will result from the package of data provided to Europeana (Fig. 1):

- the “provided cultural heritage object” itself (a painting, a movie, a music score, a book) (**edm:ProvidedCHO**);
- one or more accessible digital representations of this object, some of which will be used as previews (the digital picture of the painting) (**edm:WebResource**);
- an aggregation to represent the result of this provider’s activity. (**ore:Aggregation**).

The first two allow one to capture the distinction between “works”, which are expected to be the focus of the users’ interest, and their digital representations, which are the elements manipulated in the information systems like Europeana. The third, following the ORE approach, demonstrates that the provided object, together with the digital representations from one Europeana data provider can be regarded as one logical whole.

Aggregations enable one to capture a description of the “digital environment” of an object submitted to Europeana, and attaching descriptive information to the various resources that take part in this environment. This mechanism remains agnostic with regard to which descriptive data should be provided. EDM therefore includes a set of “descriptive” and “contextual” properties that capture the different features of a resource, as well as relating it to the other entities in its context.

Among the possible approaches for descriptive metadata, one can distinguish “*object-centric*” and “*event-centric*” approaches. EDM provides constructs that allow one to represent metadata to follow either approach. There are also classes in EDM that allow one to capture rich data. This section deals with these in order of complexity starting with the object centric approach, then looking at enriching this data with contextual classes and finally looking at the more complex event-centric approach.

The **ObjectCentric** model focuses on the object described: information comes in the form of statements that provide a direct linking between the described object and its features. They are simple strings or more complex resources denoting entities from the real world. Most metadata practices making use of the Dublin Core metadata set [DC] can be seen as an application of such an approach.

To support the modelling of such semantic enrichment and to support further enrichment, EDM features a number of classes devoted to the representation of “**contextual**” entities:

- edm: Agent: to be used for representing persons or organizations;
- edm: Event: for events;
- edm: Place: for spatial entities;
- edm: TimeSpan: for time periods or dates;
- skos: Concept: for all entities from knowledge organization systems like thesauri, classification schemes (including some place gazetteers or person authority files).

Event-centric approaches consider that descriptions of objects should focus on characterizing the various events in which objects have been involved. The idea is that it will lead to establishing richer networks of entities - by representing the events that constitute an object’s history - than with the object-centric approach. This approach underlies models such CIDOC-CRM³ and may suit the data of some (of course not all) Europeana providers.

It is not the aim of EDM to capture the full complexity of a model like CIDOC-CRM. Nor can it capture the full diversity of the all object-centric models. Rather, EDM provides a small set of properties and classes to which more specialized constructs can be “attached”.

The “*EDM mapping guideline*”⁴ report gives guidance for providers wanting to map their data to EDM. It contains definitions of the properties, information about the data types that can be used as values and the obligation level of each property. It also has an example of original data, the same converted to EDM and diagrams showing the distribution of the properties amongst the classes. “*The EDM object templates*”⁵ shows which properties apply to which class and states the data type and the obligation of the values.

³ <http://cidoc.ics.forth.gr/>

⁴ “*Europeana Data Model. Mapping Guidelines v.1.0.1*” (24/02/2012), <http://pro.europeana.eu/documents/900548/ea68f42d-32f6-4900-91e9-ef18006d652e>

⁵ <http://europeanalabs.eu/wiki/EDMObjectTemplatesProviders>

CARARE Schema

CARARE⁶ is a Best Practice Network designed to involve the network of heritage organisations, archaeological museums, research institutions and specialist digital archives in making the digital content that they hold available to Europeana, establishing a sustainable aggregation service for archaeology and architecture resources and, finally, enabling the integration of 3D and Virtual Reality content in Europeana.

As the archaeological and architecture heritage domain, belonging to a broad and diverse set of heritage organizations across Europe, was described according to different standards and management procedures, one of the CARARE's main objectives was to ensure an interoperability between the native metadata, held by the heritage organisations, and the metadata used by Europeana. For the integration of digital resources representing such heterogeneous assets, one of the most relevant task of CARARE was to create a metadata schema able to map the existing original metadata into a common output schema⁷.

Conceptually the CARARE record is focussed on an heritage asset and its relations to digital resources, activities and to collection information. 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). The cardinality of themes and elements has been specified to enable the harvesting of the real data actually present in the CARARE content providers' datasets and to meet Europeana's requirements.

Mapping CARARE on EDM

The digital content submitted by the project has been described using the EDM 5.2⁸. As EDM is a conceptual model the CARARE elements have been mapped directly to EDM Classes or to particular EDM paths. On the basis of this approach a monument has been considered an instance of the

⁶ CARARE is funded by the European Commission's ICT Policy Support Programme - contract n° 250445. The project started on 1st February 2010. www.carare.eu

⁷“Carare Metadata schema outline v.1.1” (November 2011), <http://www.carare.eu/eng/content/download/4277/35150/file/CARARE%20metadata%20schema%20outline%20v1.1.pdf>

⁸ Deliverable D2.2.3_3.4 “EDM-CARARE metadata mapping” (January 2011), http://www.carare.eu/eng/Media/Files/D2.2.3_3.4-EDM-CARARE-metadata-mapping

Edm:ProvidedCHO class and the features that refers to its nature have been grouped under the this EDM entity (Fig. 2).

As already described, CARARE first established a domain specific metadata schema based on the existing standards from the archaeology and architecture and then further work was carried out to create a mapping from the CARARE schema to EDM. Using EDM, CARARE modelled a network of connections between the heritage assets themselves and other resources that are related to the assets. These resources describe either specific real objects with their digital representations or born-digital objects. In addition, it uses the EDM contextual classes to model entities such as places and concepts.

CARARE's initial mapping to EDM focused on the three core classes implemented by Europeana in 2011⁹. Some of the elements of the core classes were not included in the initial implementation by Europeana but have since been added, e.g. in **edm:WebResource**. Europeana also implemented some of the contextual classes during 2011-12.

The CARARE mapping uses EDM classes to describe different types of information about:

- Heritage assets, such as monuments, buildings or other real world objects, identified by a set of particular characteristics that refer to their identity, location, related events, etc. Information carried by a heritage asset includes: textual metadata (such as title, etc.), thumbnails and other digital objects;
- Real world cultural objects, with their digital representations that provide other drawings and sources of information about the heritage asset (historic photographs, publications, archive materials etc.);
- Born-digital resources related to these objects, such as 3D models.

EDM requires a unique identifier for each class. Since unique identifiers could not be guaranteed in the original data, CARARE creates new identifiers for each resource based on the local identifier provided by the project partners.

In EDM, each **ProvidedCHO** gives rise to an aggregation class, whose role is to bundle a **ProvidedCHO** with **WebResources** (digital

⁹ "EDM Case Study: CARARE and EDM" <http://pro.europeana.eu/documents/900548/8303c3d2-10b8-4f36-9db5-6a3dd06bd400>

representations of the provided objects made available on the Web). CARARE creates unique identifiers for each **ore:Aggregation resource**.

These aggregation identifiers are web-enabled, in the sense that they redirect to a landing page that CARARE creates for each object. The landing page provides a unique identifier that can be used in the data, but also visually aggregates the Heritage asset with its representations, collection information and information about related activities.

In terms of EDM representation, the different information sources presented on the landing page of an archaeological asset give rise to different **ProvidedCHOs**. Each print, map, or book about an archaeological place counts as a separate object provided to Europeana. These new identifiers assigned to each object make it possible to create an explicit link between the heritage asset and a related object. For instance, the relation between a heritage asset and a document describing it is expressed using the property **edm:isRepresentationOf** with the URIs of the corresponding resources (Fig. 3).

The datasets aggregated by CARARE present a lot of differences in terms of data granularity. When converting this data to EDM, CARARE had to make decisions that are specific to each dataset.

For each cultural heritage object (a monument, a building or other real world object), CARARE provides relations to multiple digital resources. EDM allows these different resources to be connected to each other, and also permits detailed description of these resources. One of the requirements of EDM is the separation of information related to a cultural heritage object from the information describing the digital representation of this object. This distinction is particularly important when dealing with rights metadata: an object and its digital representations might have different, or even contradictory, rights statements which determine the conditions for re-use of the content. In addition to the rights information provided in the **ore:Aggregation**, CARARE provides rights information specific to these digital resources.

The spatial dimension is an important concept in the archaeology and architecture heritage domain. EDM allows the representation and the description of entities such as places, by a specific class: **edm:Place**. CARARE uses the **edm:Place** class to describe information related to a specific place separately from the **ProvidedCHO**. The relation between the **ProvidedCHO** and its related place is created via the property

dcterms:spatial. Each place has spatial coordinates and a label, sometimes provided in different languages. The inclusion of spatial coordinates in the data enables archaeology and architecture data sets to be included in geo-portals alongside other datasets used in planning, development control, tourism and other map-based services. In the Europeana context, the spatial information is used to enrich other datasets and to provide new features such as map browsing.

CARARE has completed a mapping to the **edm:Agent** class, but this has not yet been implemented in the CARARE to EDM transformation. In the current implementation of **edm:Agents** are related to **providedCHOs** through **dc:contributor**; **dc:creator** or **dc:publisher**.

In the current implementation of EDM, Concepts may be linked to **providedCHOs** through **dc:subject**. CARARE has completed a mapping to the **skos:Concept** class, although this has not yet been implemented in the CARARE to EDM transformation. This is mostly because the project has not identified any commonly used standardized vocabularies.

The CARARE repository archives data provided by the partners. Using this it is possible to retrieve records encoded according the CARARE schema or in EDM. Content partners mapped their datasets to the CARARE schema using the MINT tool. An XSLT file prepared by the project performed the transformation from CARARE data to EDM schema.

Provenance and Paradata

Provenance¹⁰

According to the W3C Provenance Incubator Group the Provenance of a resource “*is a record that describes entities and processes involved in producing and delivering or otherwise influencing that resource*”¹¹. In the Cultural Heritage domain scientific data are based on results from observation, in particular on the measurements by devices creating digital output. The systematic large-scale production of digital scientific objects, the diversity of the processes involved and the complexity of describing

¹⁰ The paragraph is based on: Doerr, M. Theodoridou, M., CRM_{dig}: A generic digital provenance model for scientific observation, *Proceedings of TaPP 2011: 3rd USENIX Workshop on the Theory and Practice of Provenance, Heraklion, Greece* (June 2011). http://static.usenix.org/events/tapp11/tech/final_files/Doerr.pdf

¹¹ W3C Provenance Incubator Group. <http://www.w3.org/2005/Incubator/prov/wiki/>

historical relationships amongst them, requires the necessity for an innovative knowledge management system. In order to capture and handle all the semantic information linked to the monitoring, management and documentation of the origins and derivation of the digital resources, recently CRM_{dig} has been implemented as an extension of the event-model CIDOC-CRM¹² ontology.

CIDOC-CRM was chosen, as the core conceptual schema, because it describes the archeological findings as it is observed or documented by the archaeologists, historians and museum experts of all the disciplines. It is an event-centric core model, implemented, among other forms, in RDFS and OWL. Instances of the CIDOC-CRM model can be merged to huge meaningful networks of knowledge about historical facts and contextual relationships. The use of the CRM and extensions of it enables an easy integration of provenance data with descriptions of the observed reality and integrated reasoning.

Scope of the CRM_{dig} is to describe all the processes starting at the level of human activities or actions, which in turn, among others, initiate "machine events" on devices and computers and form a connected graph through the data and things involved in multiple events, in roles such as input and output. The relevant context of these actions comprise descriptions of objects, people, places, times which in turn may be related to other things.

Recently in the context of the European Project 3D-COFORM¹³, addressed to a large-scale production of 3D objects for scientific and cultural use, the CRM_{dig} schema has been tested, revisited and implemented in RDFS¹⁴.

CRM_{dig} is able to model the physical circumstances of scientific observation resulting in digital data. The ontology is particularly appropriate to describe typical workflows (acquisition, processing, synthesis, presentation) creating a complex semantic network of relationships and to support complex queries which can be resolved by following deep data paths of direct or inferred relationships in the semantic network. Depending on the quality of the required reasoning, more specializations may be

¹² <http://www.cidoc-crm.org>

¹³ <http://www.3d-coform.eu>

¹⁴ http://www.ics.forth.gr/isl/rdfs/3D-COFORM_CRMdig.rdfs

introduced. It contains the constructs at the level of OPM and other models, and even more, since it is integrated in ISO21127, which allows one to connect the Provenance view with other parts of reality.

By considering the objectives of 3D-ICONS the CRMdig model has been chosen because it allows a simple and clear description of the processes carried out to digitize and render a 3D model rather a definition of the provenance in terms of metadata creation that is the key point of the DCMI's Provenance Task Group¹⁵.

Paradata¹⁶

Paradata have been defined as information about human processes of understanding and interpretation of data objects. Examples of paradata include descriptions stored within a structured dataset of how evidence was used to interpret an artefact, or a comment on methodological premises within a research publication. It is closely related, and somewhat different in emphasis, to “contextual metadata”, which tend to communicate interpretations of an artefact or collection, rather than the process through which one or more artefacts were processed or interpreted.

Numerous scientific initiatives have underlined the importance in ensuring both that the computer-based visualisation methods are applied with scholarly rigour, and that the outcomes of research including computer-based visualisation should accurately convey to users the status of the knowledge that they represent, such as distinctions between evidence and hypothesis, and between different levels of probability. A set of principles is therefore needed to ensure that digital heritage visualisation be, and it is seen to be, at least as intellectually and technically rigorous as longer established cultural heritage research and communication methods. Such principles, reflecting the distinctive properties of computer-based visualisation technologies and methods, must be made explicit like any other information concerning the cultural object. This approach allows one

¹⁵ <http://dublincore.org/groups/provenance/>

¹⁶ The paragraph is based on: Denard, H., A New Introduction to the London Charter Bentkowska-Kafel, A., Baker, D. & Denard, H. (eds.) Paradata and Transparency in Virtual Heritage, *Digital Research in the Arts and Humanities Series* (Ashgate, 2012), 57-71; http://www.londoncharter.org/fileadmin/templates/main/docs/ch6_denard.pdf

to describe the features of the physical object and its digital replicas also as far as the digitization process is concerned. The London Charter¹⁷ is an international initiative seeking to establish principles for the use of computer-based visualisation methods and outcomes in the research and communication of cultural heritage. The Charter defines principles for the use of computer-based visualisation methods in relation to intellectual integrity, reliability, documentation, sustainability and access, particularly in order to promote intellectual and technical rigour in the digital heritage visualisation, and to ensure that computer-based visualisation processes and outcomes could be properly understood and evaluated by users.

In doing so, the Charter aims to enhance the rigour with which the computer-based visualisation methods and outcomes are used and evaluated in the heritage contexts, thereby promoting understanding and recognition of such methods and outcomes. The Charter does not seek to prescribe specific aims or methods, and rather establishes those broad principles for the use, in research and communication of the cultural heritage, of the computer-based visualisation upon which the intellectual integrity of such methods and outcomes depend.

For the objectives of 3D-ICONS particularly relevant is the principle 4.6 on the Documentation Process (Paradata). The Charter establishes that *“Documentation of the evaluative, analytical, deductive, interpretative and creative decisions made in the course of computer-based visualisation should be disseminated in such a way that the relationship between research sources, implicit knowledge, explicit reasoning, and visualisation-based outcomes can be understood”*.

CARARE 2.0 Schema

As one of the main 3D-ICONS goals was to develop a metadata schema able of capturing all the semantic present in the digitation processes (*provenance*) and in understanding and interpreting data objects (*paradata*), in the first year the project has updated the CARARE schema by adding classes or entities and properties to make the original schema compliant to the 3D-ICONS requirements (Fig. 4). As CARARE worked

¹⁷ London Charter for the Use of 3D Visualisation in the Research and Communication of Cultural Heritage", was circulated and was adopted as the first official draft on June 2006; <http://www.londoncharter.org/>

like an intermediate schema between native metadata and EDM, another step was to check the compatibility between CARARE and the latest advancements of EDM.

Thanks to the recent developments of integration between CARARE and EDM and to the publication of object templates of EDM¹⁸, the updating of CARARE resulted in a simplification. The last OWL version of EDM¹⁹ has been aligned to CIDOC-CRM Core Classes and some properties of CIDOC-CRM have been reused in EDM allowing a more simple integration of CRM_{dig} into EDM.

The areas of interests of provenance and paradata are innovative for CARARE schema, but it is possible to add this information without substantially changing the original schema or the mapping on EDM.

In CARARE the Information Resource designed to hold metadata about an event is the **Activity**. In this description of different digitization processes it is important to distinguish the different types of events. For this reason the schema proposes the extension of the definition of the Event Type attribute for Activity. In Event Type is possible to record the general purpose and specific purpose of the event as well. The changes proposed for this theme are:

- **EventType** - general classification of the type of event or activity which took place, e.g. survey, archaeological excavation, digitization, rebuilding. Use of a controlled vocabulary is recommended;
- **Methods** - the methods used in this specific activity, e.g. open area excavation, sample survey, augering, boring, stratigraphic, restoration, conservation, re-pointing, photogrammetric survey etc. Use of a controlled vocabulary is recommended.

Other information that should be provided according to the principles of the London Charter have the scope to define aims and objectives of the 3D data-capture. The changes proposed for this theme consists of two new attributes for the theme Activity:

¹⁸ <http://europeanalabs.eu/wiki/EDMObjectTemplatesProviders>

¹⁹ <http://europeanalabs.eu/svn/europeana/trunk/corelib/corelib-solr-definitions/src/main/resources/eu/rdf/EDM-v524-120820.owl> (December 2012).

- **Had General Purpose** (source = CIDOC CRM) - this is a free text description of the general goal or purpose of an Activity. For example this could include practicing, preparing, monitoring, researching, designing, testing etc;
- **Had Specific Purpose** (source = CIDOC CRM) - a free text note describing the specific goal or purpose of this activity. For example, carrying out 3D data acquisition, restoration of a part of a building, completing a survey, constructing a building, etc.

To register digital machines and software used for the digitization process and formal derivation in CARARE V1.0 the Activity theme included “Material and Techniques Used”. The schema specified this element as a simple string and recommended the use of a controlled vocabulary. No distinction was made between digital machines and software (concerning different processes) it was necessary to propose a particular mapping to EDM to avoid possible misunderstandings. Each technique and/or material has to be compulsory described in separate field: for example “material and techniques=laserscanner Faro Cam2”; and “material and techniques=Meshlab”. If one register “material and techniques=laserscanner Faro Cam2; Meshlab” it was not possible to have a correct mapping with EDM. So it is mandatory to assign a specific term for each activity (digitization or processing) that the object has taken part in. Consequently, two separate elements were defined in the CARARE v2.0 schema:

- **Techniques** - the techniques used in this specific activity. Use of a controlled vocabulary is recommended;
- **Materials** - the materials used during the event method. Use of a controlled vocabulary is recommended.

In order to define the digitization process adopted to create the 3D final model the schema adds some relation addressed to explain the relations between the Heritage Asset, Digital Resource and the Activity related to the digitization processes. The properties reuse those from CIDOC-CRM and in particular from CRM_{dig}.

The relations proposed are:

- **Was_digitized_by** - this is the relation between a Heritage Asset and an Activity in which it was digitized. (It is a specialization of Was Present At). Give the id number of the target record or a URI;
- **Has Created** (source = CRMdig) this is the relation between an Activity and a digital resource or digital file that it created; includes raw data files, processed data files and final models published online. Give the id of the target record, the file-name or a URI;
- **Consists of** (source = CRMdig) this is a repeating group of elements which allows the specific activity (or activities) that took place during the overall Event to be described;
- **Created Derivative** - this is the relation which defines the reuse of a Digital Resource to create derivatives, e.g. during the different processing phases of digitization. It is a specialization of Is Derivative Of. Give the id number of the target record or a URI.

CARARE Templates and Mapping to EDM

There has been an updating of the mapping from CARARE to EDM to include the relations introduced in CARARE (Figs. 5 and 6).

- **Was_digitized_by** is a specialization of edm:wasPresentAt and associates the class ProvidedCHO with Event:

CARARE	HeritageAsset	was_digitized_by	Activity
EDM	edm:ProvidedCHO	was_digitized_by	edm:Event

- **Has_created** is a specialization of edm:wasPresentAt and associates an Event with WebResource;

CARARE	Activity	has_created	DigitalResource
EDM	edm:Event	has_created	edm:WebResource

- **Consists_of** is a specialization of dcterms:hasPart and defines a relation between two Events;

CARARE	Activity	consists_of	Activity
EDM	edm:Event	consists_of	edm:Event

- **Had_general_purpose** is a specialization of dc:description;

CARARE	Activity	had_general_purpose	literal
EDM	edm:Event	had_general_purpose	Literal / reference

- **Had_specific_purpose** is specialization of edm:isRelatedTo and links two Events;

CARARE	Activity	had_specific_purpose	Activity
EDM	edm:Event	had_specific_purpose	edm:Event

- **Created_derivative** associates an Event with a WebResource;

CARARE	Activity	created_derivative	DigitalResource
EDM	edm:Event	created_derivative	edm:WebResource

In order to capture and record the devices and the software utilized for the digitisation process it needs extend the proposed schema using other specific elements from CRM_{dig}. In particular the schema proposes other two relations both as specialization of Edm:was_Present_At:

- **Used_software_or_firmware** - associates an Event with an InformationResource²⁰;

CARARE	Activity	Materials	Literal
EDM	edm:Event	Used_software_or_firmware	Edm:InformationResource

- **Happened_on_device** - associates an Event with a NonInformationResource;

CARARE	Activity	Materials	Literal
EDM	edm:Event	Happened_on_device	Edm:NonInformationResource

²⁰ For *Information resource* and *non information resources* elements, that are not still implemented in the Object_Template, see the *EDM Data Model elements, Version 5.2.3 (24/02/2012) Europeana v 1.0*.

Furthermore to define digital machines and software it is possible to use the following concepts (source CRM_{dig}) mapped on the corresponding Edm:entities:

- DigitalDevice = Edm:NonInformationResource (= CARARE:Materials);
- Software = Edm:InformationResource (= CARARE:Materials).

In order to map correctly CARARE:Materials the schema proposes the following paths:

- If the value in CARARE for Activity/Event_type is = DigitizationProcess then the value for “materials” in CARARE corresponds to an Edm:NonInformationResource = DigitalDevice and the relation linking the event to the NonInformationResource is edm:WasPresentAt = Happened_on_device;
- If the value of Activity/Event_type is = SoftwareExecution or Formal Derivation then the value of materials is an Edm:InformationResource = Software and the relation linking the event to the edm:InformationResource is edm:WasPresentAt = Used_software_or_firmware.

Conclusions

The report on Metadata and Thesauri, addressed the updating of the original CARARE Schema. The updated schema is a good compromise between the need to store a complete track of the digitization process of the digital object and an easy way to standardized the metadata for 3D models. The extension of the previous CARARE schema covering paradata and provenance concepts assure a simple mapping among different existing standard. Hopefully CARARE 2.0 will foster the adoption by European Institutions of a clearer approach to describing the features of the cultural object, the techniques and the methodologies chosen for the digitization and the motivation behind the creation of the digital object. Complete knowledge of the digital resource will allow for a more efficient reuse of the archive and increase the usability of the resources available on-line. Furthermore it will be easier to compare models, their complexity, any eventual innovation in their creation and their reliability. The availability within Europeana of 3D models of architecture or archaeological

monuments from the 3D ICONS project will show best practices and guidelines useful for new projects and 3D data collections. The metadata provided to Europeana will be published under the terms of the Creative Commons CC0 1.0 Universal Public Domain Dedication and can therefore be re-used by third parties without any restrictions.

Europeana recently launched the portal data.europeana.eu²¹ to make metadata openly available to public and private sectors alike so they can use it to develop innovative applications and create new web services and portals,. So far only a few collections have been published according to the Linked Open Data (LOD) protocol. This is a data publishing technique that uses common web technologies to connect related data and make them accessible on the Web. The data are represented in EDM and the described resources are addressable and de-referencable by their URIs. Links between Europeana resources and other resources in the Linked Data Web enable the discovery of semantically related resources.

In the future all Europeana metadata will be published as Linked Open Data thus enabling data providers to benefit from Europeana's efforts to link them to semantically related resources on the Web.

²¹ <http://pro.europeana.eu/linked-open-data>

REFERENCES

D'Andrea, A. (2012) Integrating Architectural and Archaeological 3D-Models into Europeana, *Newsletter di Archeologia CISA*, 3, 87-109. Napoli.

(all web-sites have been visited on 30/01/2013)

3D-COFORM = <http://www.3d-coform.eu>

CARARE = www.carare.eu

CARARE REPOSITORY = <http://store.carare.eu>

CARARE Metadata schema outline v.1.1” (November 2011)

<http://www.carare.eu/eng/content/download/4277/35150/file/CARARE%20metadata%20schema%20outline%20v1.1.pdf>

CIDOC-CRM RDFS 5.0.4 = http://www.cidoc-crm.org/rdfs/cidoc_crm_v5.0.4_official_release.rdfs

CRMdig = Doerr, M. Theodoridou, M. “*CRMdig: A generic digital provenance model for scientific observation*”, Proceedings of TaPP 2011: 3rd USENIX Workshop on the Theory and Practice of Provenance, Heraklion, Greece (June 2011). <http://www.usenix.org/events/tapp11>

CRMdig RDFS 2.5 = http://www.ics.forth.gr/isl/rdfs/3D-COFORM_CRMdig.rdfs

DCMI Provenance Task Group = <http://dublincore.org/groups/provenance/>

EDM-CARARE metadata mapping” (January 2011)

http://www.carare.eu/eng/Media/Files/D2.2.3_3.4-EDM-CARARE-metadata-mapping

EDM Case Study: CARARE and EDM

<http://pro.europeana.eu/documents/900548/8303c3d2-10b8-4f36-9db5-6a3dd06bd400>

EDM Data Model elements, Version 5.2.3 (24/02/2012) Europeana v 1.0

<http://pro.europeana.eu/documents/900548/bb6b51df-ad11-4a78-8d8a-44cc41810f22>

EDM Linked Open Data = <http://pro.europeana.eu/linked-open-data>

Haslhofer, B., Isaac, A., “*data.europeana.eu - The Europeana Linked Open Data Pilot*”, DCMI Global Meetings & Conferences, DC-2011, The Hague, <http://dcevents.dublincore.org/IntConf/dc-2011/paper/view/55/14>

Isaac, A., Clayphan, R., Haslhofer, B., "Europeana: Moving to Linked Open Data", *ISQ Information Standards Quarterly* Spring/Summer 2012, volume 24, issues no. 2/3, http://www.niso.org/apps/group_public/download.php/9407/IP_Isaac-et-al_Europeana_isqv24no2-3.pdf

EDM Mapping Guidelines v.1.0.1 (24/02/2012)

<http://pro.europeana.eu/documents/900548/ea68f42d-32f6-4900-91e9-ef18006d652e>

EDM Object Templates = <http://europeanalabs.eu/wiki/EDMObjectTemplatesProviders>

EDM Primer (26/10/2011) Europeana v1.0

<http://pro.europeana.eu/documents/900548/770bdb58-c60e-4beb-a687-874639312ba5>

EDM RDFS 5.24-120820 = <http://europeanalabs.eu/svn/europeana/trunk/corelib/corelib-solr-definitions/src/main/resources/eu/rdf/EDM-v524-120820.owl>

PARADATA = Denard, H., A New Introduction to the London Charter, Bentkowska-Kafel, A., Baker, D. & Denard, H. (eds.) *Paradata and Transparency in Virtual Heritage, Digital Research in the Arts and Humanities Series* (Ashgate, 2012), 57-71. http://www.londoncharter.org/fileadmin/templates/main/docs/ch6_denard.pdf

W3C PROVENANCE = <http://www.w3.org/2005/Incubator/prov/wiki/>

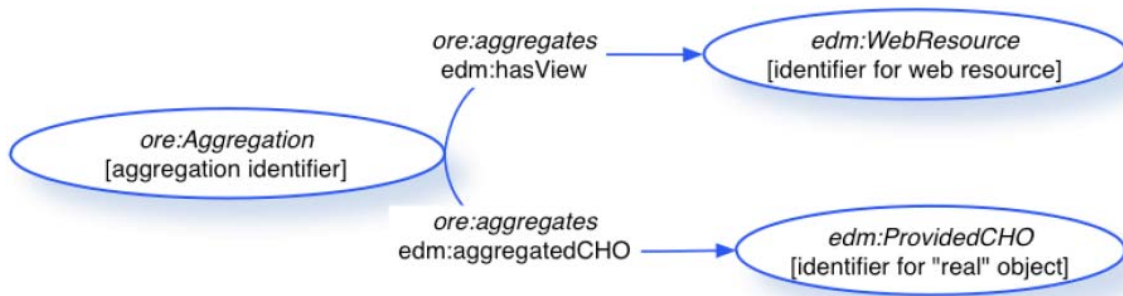


Fig. 1 - Visualization of the three core EDM classes for data providers

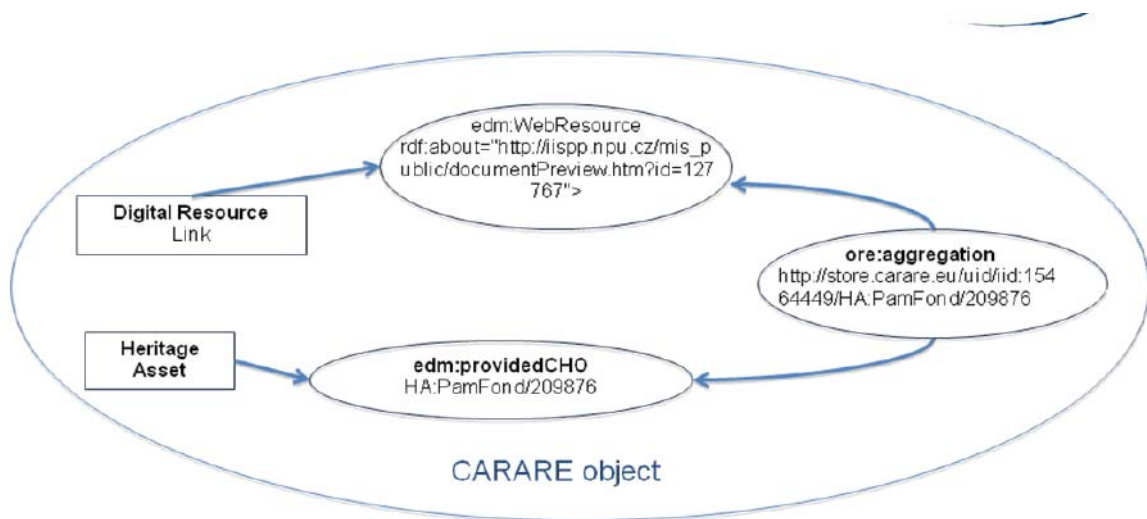


Fig. 2 - Mapping of CARARE on the three core EDM classes

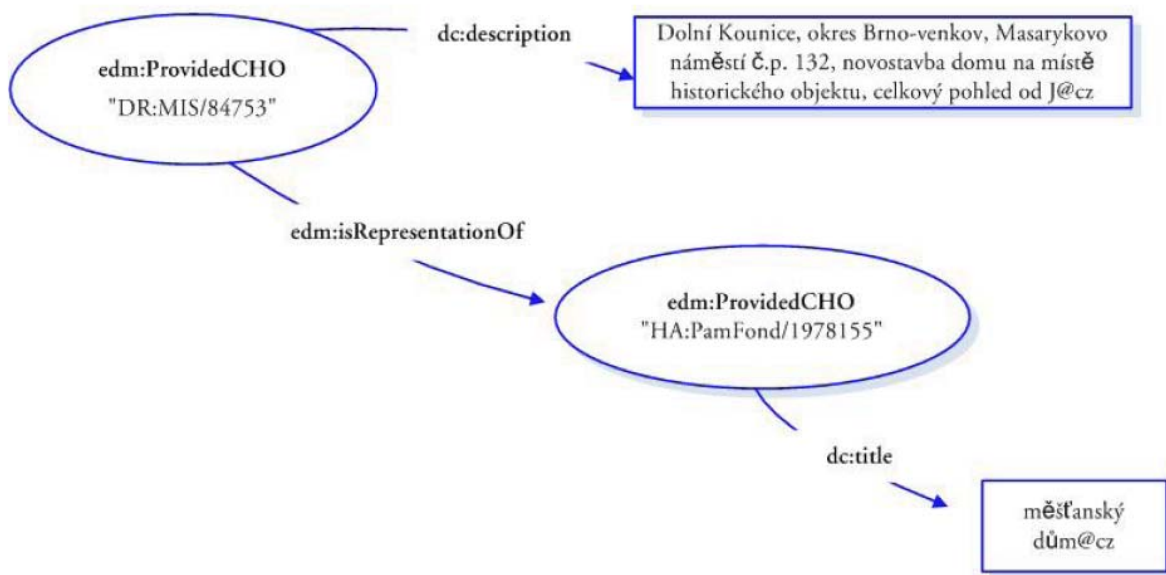


Fig. 3 - The relationship between Digital Resource (DR) and Heritage Asset (HA) mapped on the core EDM class edm:ProvidedCHO

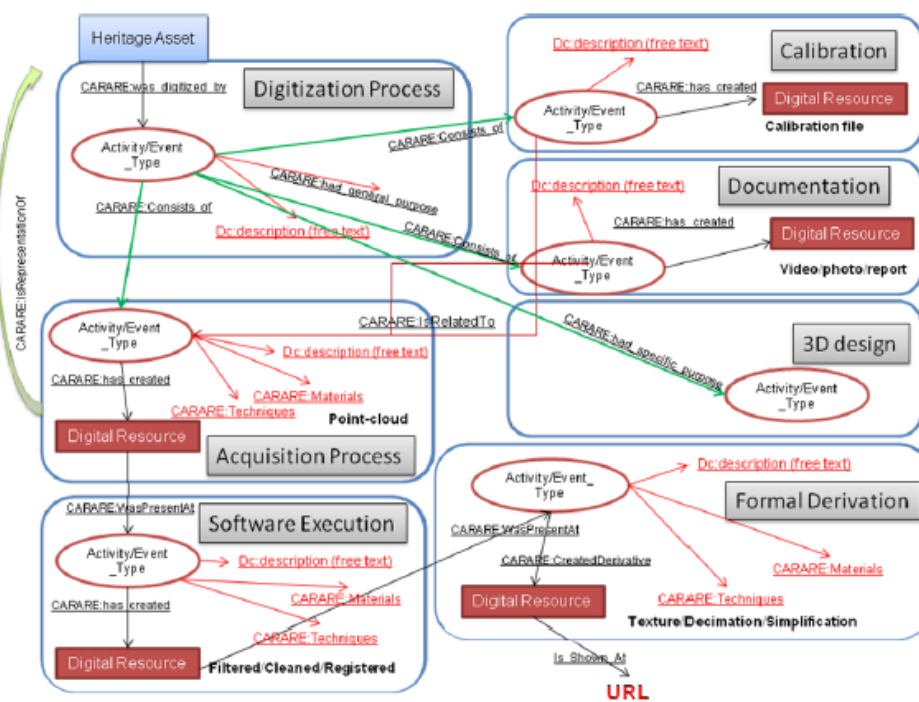


Fig. 4 - The 3D digitization workflow in the CARARE 2.0 Schema

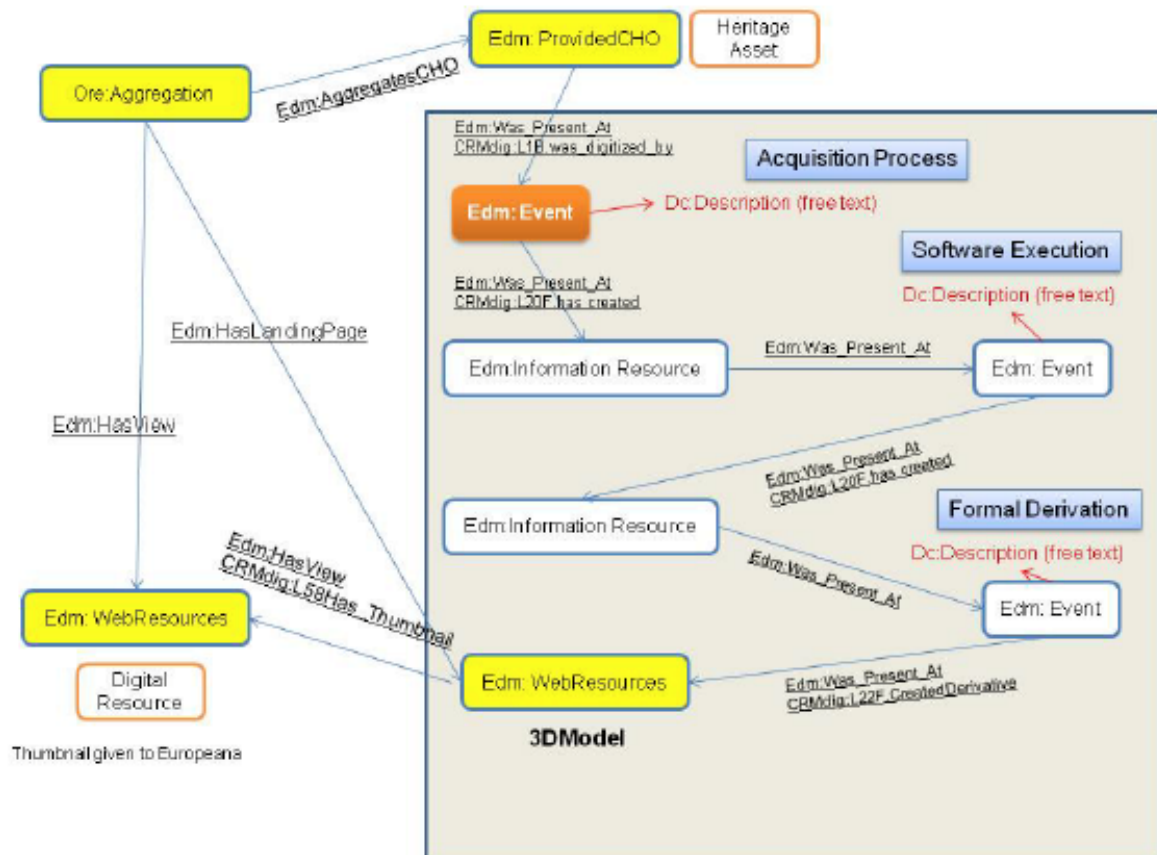


Fig. 5 - CARARE 2.0 mapped on EDM Schema: from data-acquisition process to the final 3D model on the web

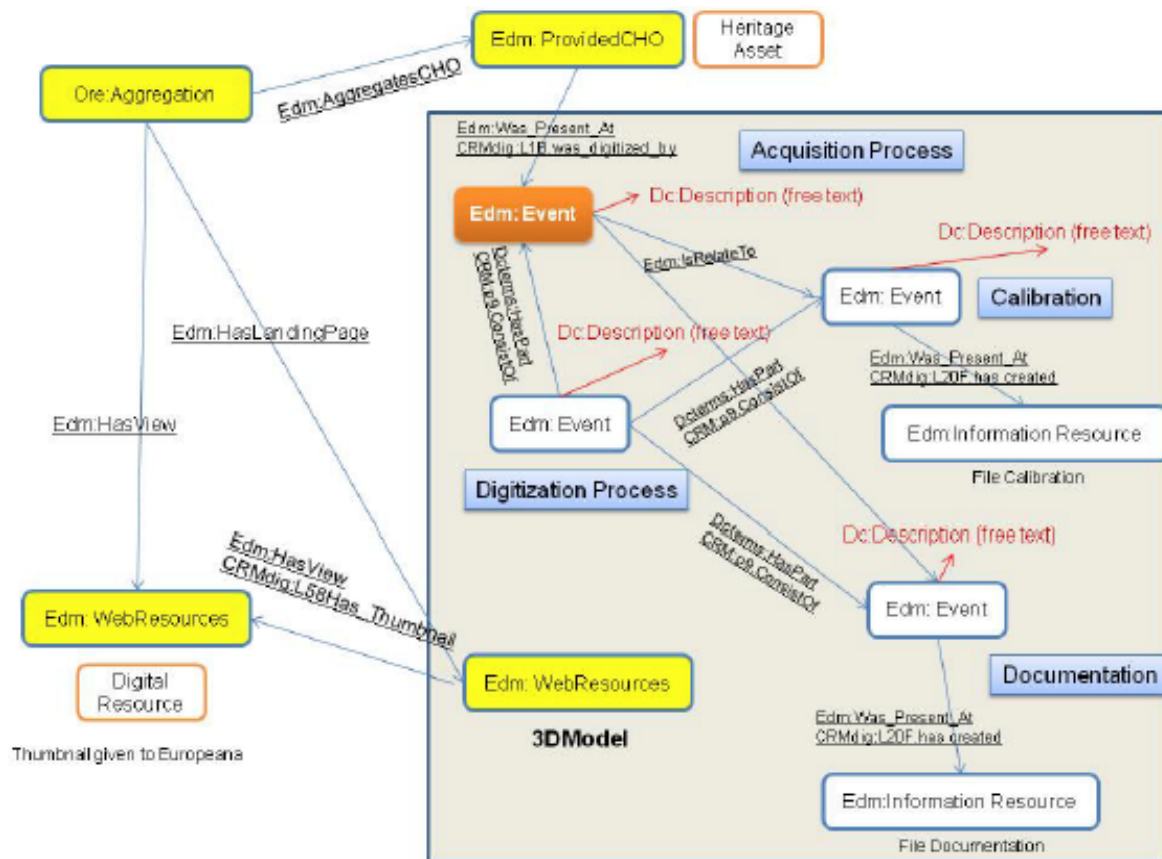


Fig. 6 - CARARE 2.0 mapped on EDM Schema: the digitization process