Mediterranean Archaeology and Archaeometry, Vol. 16, No 5,(2016), pp. 53-61

DOI: 10.5281/zenodo.204966

METADATA FOR THE LEARNING OBJECTS THAT CONTAIN CULTURAL OBJECTS

Stella Markantonatou*, Panagiotis Minos, Katerina Tzortzi, George Pavlidis

Institute for Language and Speech Processing/"Athena" RIC

Received: 05/12/2015
Accepted: 15/07/2016
Corresponding author: Stella Markantonatou(marks@ilsp.athena-innovation.gr)

ABSTRACT

Standardised metadata for the development and storage of learning objects that draw on cultural content have not been provided as yet. There are well established standards for documenting learning objects, such as the IEEE/LOM and the LRE/LOM, and for documenting cultural content, such as the CIDOC-CRM and the SPECTRUM. However, these standards do not provide for standardised metadata of learning objects that exploit cultural objects and enable descriptions of the type *Alpha is a learning object made by X and contains a digital image of Joconda made by DaVinci;* nor do they provide for a solid framework for DBs that could answer to questions such as "Provide me with learning objects supporting Geography lessons about Greece that are addressed to 14-year olds and contain pictures by one of Papaloukas, Vassileiou, Tsarouhis" or "Provide me with either learning objects or digital pictures of pieces of art on harvesting wheat". Furthermore, existing large repositories of digital objects have not been designed for such hybrid objects. Thus, EUROPEANA contains only cultural content while the LRE and the Greek Φωτόδεντρο only learning objects. In the overall, it is clear that if education at a European level is meant to profit from cultural objects and cultural object repositories according to EU plans (Eurydiki 2009), the process of developing and storing learning objects that contain cultural content should be strengthened with standardised documentation and storage tools. We report here on the development of an ontology that is the unification of LRE/LOM (the standard for learning objects) and CIDOC-CRM (the standard for cultural objects). We also report on the development of a database that conforms to the unification of LRE and CIDOC ontologies.

KEYWORDS: metadata, standardization, learning object, cultural content, LRE, LOM, CIDOC-CRM, ontology

1. INTRODUCTION

Standardised metadata for the development and storage of learning objects (LOs) that draw on cultural content (CLOs from now on) have not been provided as yet. There are well established standards for documenting LOs, such as the IEEE/LOM and the LRE/LOM, and for documenting cultural content, such as the CIDOC-CRM and the SPECTRUM. However, these standards do not provide for CLOs that would require descriptions of the type *Alpha is a LO made by X and contains a digital image of Joconda made by DaVinci* nor for a solid framework for building a DB that could answer to questions such as "Provide me with LOs supporting Geography lessons about Greece that are addressed to 14-year olds and contain pictures by one of Papaloukas, Vassileiou, Tsarouhis" or "Provide me with either LOs or digital pictures of pieces of art on harvesting wheat". Furthermore, existing large repositories of digital objects have not been designed to document the complex structure of such hybrid objects. Thus, EUROPEANA contains only cultural objects while the LRE and the Greek Φωτόδεντρο only LOs. As a result, CLOs cannot be documented, stored and reused by exploiting the international large cultural object and learning object collections.

However, cultural object based learning that requires the development of CLOs is one of the modern requirements in formal and life-long education (see for instance Lambropoulos et al. (2005), Shieber (2012), British Museum), often through serious gaming (Fletcher 2014, project GameIt). It is clear that if education at a European level is meant to profit from cultural content according to EU plans (Eurydiki 2009), the process of developing and storing CLOs should be strengthened with standardised full documentation and storage tools that will allow the developers and users of this type of LO to easily create and/or retrieve these objects. We report here on the development of an ontology that is the unification of LRE/LOM and CIDOC-CRM and enables the standardised detailed documentation of CLOs. We will call this ontology CIDOC+LRE. We also report on the development of a database that conforms with the CIDOC+LRE ontology.

In section 2, we briefly present the CIDOC-CRM and the LRE/LOM by focussing on their particular features that both necessitated and facilitated their unification. In section 3, we describe the unified ontology CIDOC+LRE in some detail. Lastly, in section 4 we present a DB that is compatible with the CIDOC+LRE ontology.

2. THE SOURCE METADATA SCHEMES

A brief description of the source ontology CIDOC-CRM and the LRE/LOM metadata scheme is given in this section.

2.1. *CIDOC Conceptual Reference Model*

CIDOC Conceptual Reference Model (CRM) (de Boeuf et al., 2015), an ISO standard[1], provides definitions and the necessary formal structure for the description of concepts and relations used in the documentation of cultural content.

The CIDOC-CRM ontology that is expressed as an object-oriented semantic model can be converted to machine-readable formats (RDF Schema, KIF, DAML+OIL, OWL etc) and can be implemented in any relational or object-oriented schema.

CIDOC-CRM contains 89 classes and 151 unique properties. A CRM class is a category of items that share one or more common traits used as identification criteria. An item that belongs to a class is called an instance of this class. Classes are possibly related to each other in two ways:

- by being the domain or the range of none, one or more properties
- by being involved in an IsA relationship.

The IsA relationship, where a subclass is a specialization of its superclass, entails that a subclass inherits all the properties declared for its superclass without exceptions (strict inheritance), in addition to having none, one or more properties of its own. A subclass can have more than one immediate superclass (multiple inheritance).

The top class of CIDOC-CRM is the E1 CRM entity with 6 subclasses as shown in Figure 1. There are two interesting features of the CIDOC-CRM ontology, namely:

- both the material and the immaterial aspects of the nature of a cultural object can be documented, as specializations of the E77 class
- processes and states concerning the live cycle of museum/cultural objects can also be documented as specializations of the E2 class

CIDOC-CRM is an ontology expressive enough to document the properties and the life-cycle of objects, such as the LOs, that are the result of human intellectual activity. With the reported work, the CIDOC-CRM ontology has been extended to accommodate LRE/LOM that is tailored to the needs of LO documentation.

[1] ISO 21127:2014

E1 CRM Entity
Superclass of:
E2 Temporal Entity
E52 Time-Span
E53 Place
E54 Dimension
E77 Persistent Item
E92 Spacetime Volume

Figure 1. The CRM top entity E1

2.2. *Learning Resource Exchange -Learning Object Metadata*

Learning Resource Exchange (LRE) is an international repository. Its functionality rests on a combination of IMS LODE Information for Learning Object eXchange specification (ILOX) with IEEE Learning Object Metadata (LOM) IEEE 1484.12.1 – 2002 Standard for Learning Object Metadata.

LOM describes LOs and resources for the education aiming at facilitating retrieval. It is an open standard with wide acceptance. However, LOM does not provide all the necessary metadata for the LOs that are stored in LRE.

In LRE, a LO is defined as anything digital used for teaching or learning (e.g., an online course, an image, an assessment). However, several versions of a learning object might exist (an English and a Greek version for instance), each one probably in different formats and in a number of copies. These properties are critical for the discovery and exchange of LOs and this is why LRE uses as a data scheme a combination of ILOX and LRE as follows (Figure 2):

- "Work" is the root element and corresponds to the abstract concept of LO
- "Expression", corresponds to the concept of version
- "Manifestation" corresponds to the concept of format
- "Item" corresponds to the concept of copies.

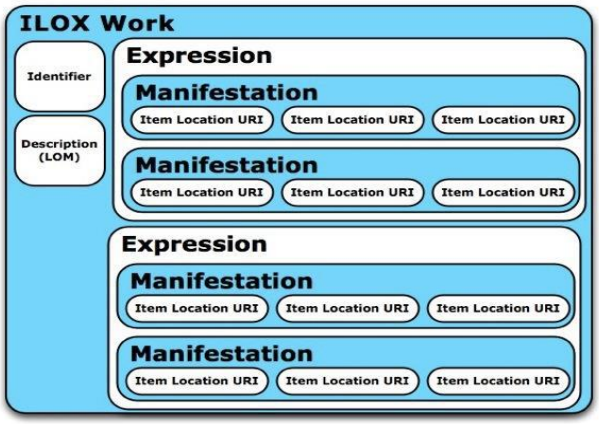

Figure 2. Overview of an LRE metadata instance with Work as root (Massart et al, 2011:12)

Furthermore, each of these classes is a construct that can be described using four types of element: an identifier, a general description, construct-specific information (if needed), and one or more concrete constructs.

Table I. Definition of the top node of the three ontologies

CIDOC+LRE	CIDOC	LRE-MAP 'WORK'
E1 CIDOC+LRE: Anything that can be said to describe a cultural object or a LO.	E1 CRM Entity: This class comprises all things in the universe of discourse of the CIDOC Conceptual Reference Model. It is an abstract model providing for three general concepts: 1. Identification by name or appellation, and in particular by a preferred identifier 2. Classification by type, allowing further refinement of the specific subclass an instance belongs to 3. Attachment of free text for the expression of anything not captured by formal properties"	LRE-MAP 'WORK': 'Work' corresponds to the abstract concept of LO"

3. THE CIDOC+LRE ONTOLOGY

We now turn to the unification of the two data schemes. Of them, only CIDOC-CRM is an ontology. The aim is to derive a new ontology that offers full interoperability between databases and repositories that are based on CIDOC-CRM and LOM (Sowa, 2009).

We proceeded in the development of the CIDOC+ LRE ontology by:

- Unifying those classes and properties of the two ontologies that could be unified.
- Introducing new classes and properties if no class or property unification was possible

The top class of the new ontology CIDOC+LRE is the class CRM Entity_LRE WORK; it is the unification of the top class E1 CRM Entity of the CIDOC ontology and the top class 'WORK' of the LRE/LOM ontology. Table 1 shows the definitions of the two classes in the respective frameworks and the definition of their unification in CIDOC+LRE. Both the unified classes are very general and their content could be described as "all what one can say about a LO or a cultural object".

3.1. *Learning Objects (LOs) in CIDOC+LRE*

CIDOC-CRM, being a standard for the description of cultural objects, does not contain LOs. A new class 'LO' is defined in the CIDOC+LRE ontology as a subclass of the class *E90 Symbolic Object* of the original CIDOC ontology. The definition of **E90** (Le Boeuf et al., 2015:40) includes the following: "**E90 Symbolic Object:** This class…includes sets of signs of any nature, which may serve to designate something, or to communicate some propositional content. An instance of E90 Symbolic Object does not depend on a specific physical carrier, which can include human memory, and it can exist on one or more carriers simultaneously." The CIDOC subclasses of E90 are: *E41 Appellation* and *E73 Information Object*. E41 includes name entities while E73 comprises immaterial structured objects, namely the subclasses *E29 Design or Procedure*, *E31 Document, E33 Linguistic Object* and *E36 Visual Item*. In CIDOC+LRE, LO has been defined as a sister of E73 and E41 because, just like E73 it has content but characteristically unlikely E73 it also has learning features.

Table 2. The new entities that were defined in the CIDOC+LRE ontology

ENTITY	IS A SUBCLASS OF	DESCRIPTION
E194 LO	E90 Symbolic Object	LO
E195 Validation	E7 Activity	LO evaluation facts
E196 Meta-metadata Entry	E7 Activity	LO Metadata Insertion Event
E197 Annotation	E7 Activity	LO Usage Annotation Event
E198 Electronic location	E44 Place Appellation	LO Electronic Address

3.2. *Unifying CIDOC with LOM*

LOM defines 9 sets of features for LOs as shown in Figure 3. To document LOs according to LOM these sets of features have be unified with CIDOC+LRE. To this purpose, a small set of new entities, one of them being LO, had to be defined in the CIDOC+LRE ontology (Table 2). In what follows we give a relatively detailed account of the unification of each LOM feature with CIDOC-CRM. We use Figure 3 as a reference.

Some general points that are not mentioned in the description of the encoding below (sections 3.2.1-3.2.9) because they occur systematically are:

- The property P2 of CIDOC-CRM has been used to introduce descriptions
- Often, LOM uses controlled vocabularies that correspond to CIDOC-CRM entities, a fact that was taken into account in CIDOC+LRE.

3.2.1. (LOM 1) General in CIDOC+LRE

No new entity was defined for the LOM category *General*. Some of its properties were unified with existing CIDOC properties while a small set of new properties was defined as follows:

(LOM 1.1) *Identifier* was unified with *P1 is identified by (identifies)* as follows for LO: *E194 LO P1 is identified by (identifies) E42 Identifier*

(LOM 1.2) *Title* was unified with *P102 has title (is title of)* as follows: *E194 LO P102 has title (is title of) E35 Title*

(LOM 1.3) *Language*: The new property *E194 LO P265 contains E33 Linguistic Object* was defined to express the fact that LOs may contain texts that in CIDOC+LRE are instances of the class *E33 Linguistic Object*

(LOM 1.4) *Description* was unified with *P3 has note* as follows for LO: *E194 LO P3 has note E62 String*

(LOM 1.5) *Keyword*: The new property *E194 LO P266 has keyword (is keyword of) E62 String* was defined

(LOM 1.6) *Coverage*: Three new properties (P267, P268, P269) were defined: *E194 LO P267 has time coverage (is time coverage of) E52 Time-Span P78*, *E194 LO P268 has place coverage (is place coverage of) E53 Place* and *E194 LO P269 jurisdiction is held by (holds jurisdiction* of) *E40 Legal Body*

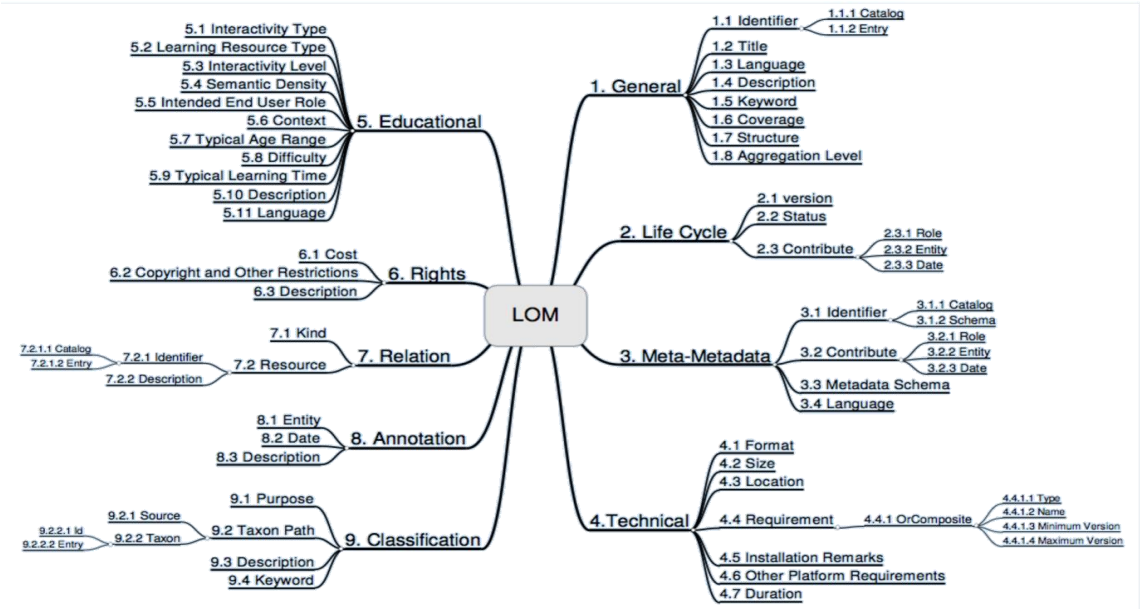

Figure 3. Schematic presentation of LOM

(LOM 1.7) *Structure* and (LOM 1.8) *Aggregation Leve*l were unified with *P2 has type (is type of)* and receive values from controlled vocabularies specified in LOM

3.2.2. (LOM 2) Life Cycle in CIDOC+LRE

No new entity was defined for the category *Life Cycle*. The features it contains were integrated as follows:

(LOM 2.1) *Version* and (LOM 2.2) *Status*: The new property *E194 LO P284 has version (is version of) E194 LO* was defined to denote the fact that a LO is at a certain condition as a result of a modification action/procedure that does not consume it (as opposed to cultural objects).

(LOM 2.3) *Contribute*: It was unified with appropriate CIDOC properties depending on the type of contribution described. More particularly:

- (LOM 2.3) *Contribute*: *E65 Creation P94 has created (was created by) E194 LO / E65 Creation P14 carried out by (performed) E39 Actor* where Actor is further specified with the following controlled vocabulary *author, graphical designer, technical implementer, script writer, instructional designer, subject matter expert*

- (LOM 2.3) *Contribute/publisher* (it has been treated along with LOM 3.2.1, see Section 3.1.4)

- (LOM 2.3) Contribute/editor: *E11 Modification P270 has edited (is edited by) E194 LO / E11 Modification P14 carried out by (performed) E39 Actor P3*

- (LOM 2.3) *Contribute/initiator*: *E63 Beginning of Existence P92 brought into existence (was brought into existence by) E194 LO / E63 Beginning of Existence P14 carried out by (performed) E39 Actor*

- (LOM 2.3) *Contribute/terminator*: *E64 End of Existence P93 took out of existence (was taken out of existence by) E194 LO / E64 End of Existence P14 carried out by (performed) E39 Actor*

- (LOM 2.3) *Contribute/validator*: A new entity E195 Validation was defined as a subclass of E7 Activity (Table 2) and the new property E195 Validation P271 has validated (is validated by) E194 LO. Validator is an instance of E9 Actor in an event that instantiates the class E195.

3.2.3. (LOM 3) Meta-metadata in CIDOC+LRE

E196 Meta-metadata Entry was defined as a subclass of *E7 Activity* (Table 2). E196 is the domain of the following properties:

P272 describes metadata record of (metadata record is described by) E194 LO

(LOM 3.2.1 creator) *P14 carried out by (performed) E39 Actor P3 has note E62 String*

(LOM 3.2.3) *P4 has time-span (is time-span of) E52 Time Span*

(LOM 3.1) *P1 is identified by (identifies) E42 Identifier*

(LOM 3.3) *P273 has schema (is schema of) E62 String* (the property corresponds to

(LOM 3.4) *P275 registration has language (is language of registration) E56 Language*

P274 is validated by (has validated) E195 Validation

3.2.4. (LOM 4) Technical in CIDOC+LRE

The fact that LO, a symbolic object, has a carrier and the properties of the carrier are modeled in CIDOC+LRE as follows:

(LOM 4.1) *Format* of the carrier: *E194 LO P128 is carried by (carries) E24 Physical Man made Thing P289 has format (is format of) E55 Type [P289.1 has format type E55 Type (format type)]*

(LOM 4.2) *Location* of the carrier: *E194 LO P128 is carried by (carries) E24 Physical Man made Thing E194 LO P53 has former or current location (is former or current location of) E53 Place P87 is identified by (identifies) E44 Place Appelation*

(LOM 4.3) *Size* of the carrier: *E194 LO P128 is carried by (carries) E24 Physical Man made Thing /E54 size of the carrier P91 has unit (is unit of) E58 Measurement Unit P90 has value E60 Number*.

3.2.5. (LOM 5) Educational in CIDOC+LRE

Some new properties but no new entity was defined for the encoding of the educational properties of LO.

(LOM 5.1) *Interactivity Type*: *E194 LO P2 has type (is type of) E55 Type [P2.1 has interactivity type E55 Type (interactivity type)]*

(LOM 5.2) Learning Resource Type: *P2 has type (is type of) E55 Type [P2.1 has learning resource type E55 Type (learning resource type)]*.

(LOM 5.3) I*nteractivity Level: E194 LO P2 has type (is type of) E55 Type [P2.1 has interactivity level E55 Type (interactivity type)]* (controlled vocabulary)

(LOM 5.4) *Semantic Density*: *E194 LO P2 has type (is type of) E55 Type [P2.1 has semantic density type E55 Type (semantic density type)]* (controlled vocabulary)

(LOM 5.5) *Intended End User Role*: *E194 LO P2 has type (is type of) E55 Type [P2.1 has intended end user role type E55 Type (user type)]* (controlled vocabulary)

(LOM 5.6) *Context*: *E194 LO P2 has type (is type of) E55 Type [P2.1 has context type E55 Type (context type)]* (controlled vocabulary)

(LOM 5.7) *Typical Age Range*: *E194 LO P276 has typical age range (is typical age range of) E62 String*

(LOM 5.8) *Difficulty*: *E194 LO P2 has type (is type of) E55 Type [P2.1 has difficulty type E55 Type (difficulty type)]* (controlled vocabulary)

(LOM 5.9) *Typical Learning Time*: *E194 LO P290 has typical learning time (is typical learning time of) E52 Time Span P3 has note E62 String*

(LOM 5.10) *Description*: *E194 LO P277 has usage description (is usage description of) E62 String*

(LOM 5.11) *Language*: *E194 LO P278 has user language (is user language of) E56 Language [P278.1 has user language type E55 Type (user language type)]* (controlled vocabulary)

3.2.6. (LOM 6) Rights in CIDOC+LRE

The CIDOC entity *E30 Right* has been unified with (LOM 6) *Rights*. The new property *P279 has cost (is the cost of) E62 String* encodes the usage conditions of *E194 LO*:

E194 LO P279 has cost (is the cost of) E62 String

E194 LO P104 is subject to (applies to) E30 Right P105 right held by (has right on) E39 Actor

Table 3. Dublin Core terms for (LOM 7) Relation

Dublin Core Terms	Description
1. is part of	A related resource in which the described resource is physically or logically included.
2. has part	A related resource that is included either physically or logically in the described resource.
3. is version of	A related resource of which the described resource is a version, edition or adaption.
4. is format of	A related resource that is substantially the same as the described resource but in another format.
5. has format	A related resource that is substantially the same as the pre-existing described resource but in another format.
6. references	A related resource that is referenced cited or otherwise points to the described resource.
7. is referenced by	A related resource that references, cites, or otherwise points to the described resource.
8. is basis for	
9. requires	A related resource that is required by the described resource to support its function, delivery or coherence.
10. is required by	A related resource that requires the described resource to support its function, delivery or coherence.

3.2.7. (LOM 7) Relation in CIDOC+LRE

(LOM 7) *Relation* specifies relations among LOs. LOM uses the Dublin Core (DC) terms for relations (Table 3). In CIDOC+LRE LOM relations among LOs were either unified with existing CIDOC-CRM relations or were incorporated as new relations as follows:

is part of (1) & *has part* (2): *E194 LO P106 is composed of (forms part of) E194 LO*

is version of (3): see section 3.2.2 (LOM 2.1)

is format of (4) & *has format* (5): see section 3.2.4 (LOM 4.1)

references (6) & *is referenced by* (7): *E194 LO P67 refers to (is referred to by) E194 LO*

is basis for (8): *E194 LO P281 is basis for (is based on) E194 LO*

requires (9) & *is required by* (10): *E194 LO P282 requires (is required by) E194 LO*

Crucially, CIDOC+LRE documents LOs that have cultural objects as their parts and vice versa. P283 represents this idea: *P283 contains (is contained by) E70 Thing*.

3.2.8. (LOM 8) Annotation in CIDOC+LRE

The new entity *E197 Annotation* was defined as a subclass of *E7 Activity* (Table 1). E7 is the domain of the following properties:

(LOM 8.1) *Entity: P14 carried out by (performed) E39 Actor*

(LOM 8.2) *Date: P4 has time-span (is time-span of) E52 Time Span P78 is identified by (identifies) E49 Time*

(LOM 8.3) *Description: P3 has note E62 String*

E194 LO P286 has classification (is classification of) E62 String

3.2.9. (LOM 9) Classification in CIDOC+LRE

No new entity was defined for (LOM 9) Classification. A set of new properties were defined as follows:

(LOM 9.1) *Purpose: E194 LO P287 has classification purpose (is classification purpose of) E55 Type*

(LOM 9.2) *Taxon Path: E194 LO P288 has taxon path (is taxon path of) E62 String*

4. THE DB

Both LOs and cultural objects can be documented in the database that was developed in accordance to the CIDOC+LRE ontology. In Figure 4 a LO about Geography is documented that uses 4 cultural objects. The first of these objects is highlighted in the Objects Window. Figure 5 shows part of the documentation of this cultural object. The database is completed with a set of controlled vocabularies. For LRE, vocabularies have been drawn from the Vocabulary Bank for Education – VBE. Several Greek controlled vocabularies, compatible with the established AAT Getty, have been developed to support documentation of cultural objects.

The GUI is an expandable, platform-independent Java application for storage and retrieval of LOs (Figure 4) and cultural objects (Figure 5) and has been created on the NetBeans Platform. It can use various Relational Database management systems (e.g. Apache Derby, MsSQL, MySQL) and makes full use of the underlying database schema. In Figure 6 the DB schema for the root entity E1 of CIDOC+LRE is given (the DB (labels, vocabularies) has been implemented for the Greek language. Future editions will be multilingual).

5. CONCLUSION

The DB is a powerful tool for the developer/user of CLOs because it enables her to exploit the contents of LO and cultural object repositories/DBs with their original metadata that document the fact that LOs and cultural objects are not identical from a conceptual point of view.

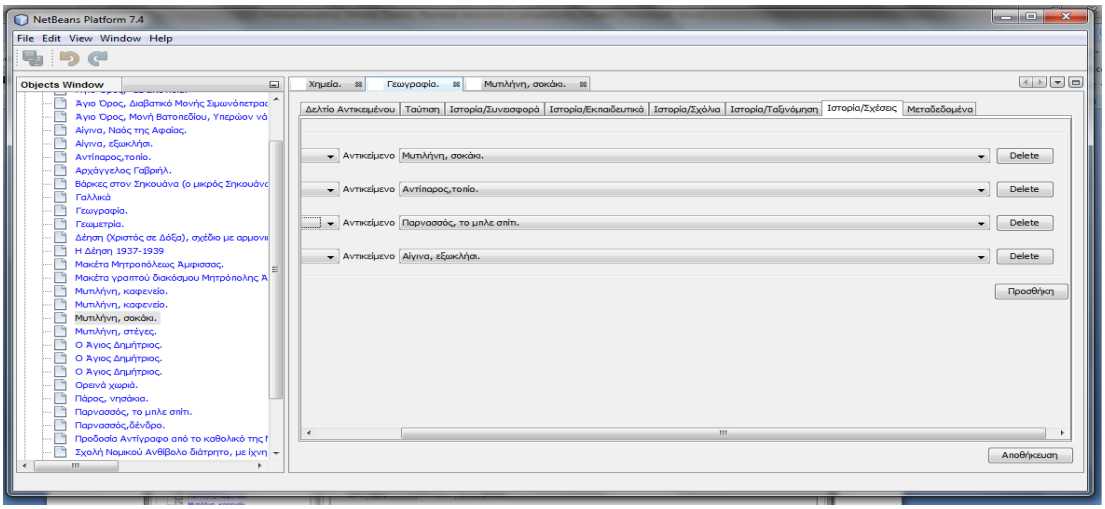

Figure 4. Documentation of Learning Object

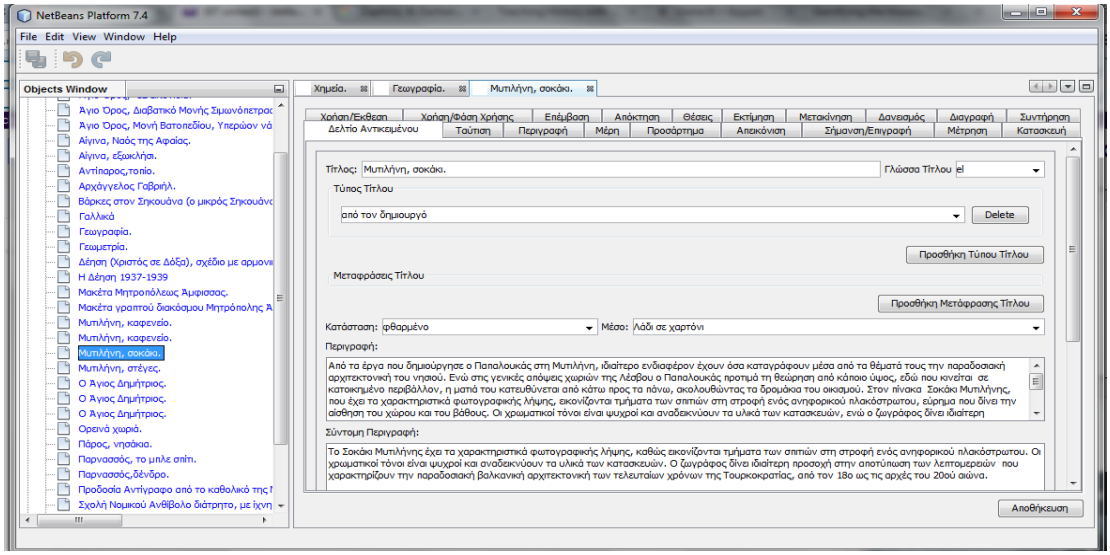

Figure 5. Documentation of Cultural Object

Figure 6. The database schema for E1 CIDOC+LRE

ACKNOWLEDGMENTS

The research reported here was conducted within the PAVET project GameIT 1/2/2014-31/6/2015 that was funded by GSRT.

REFERENCES

Φωτόδεντρο http://photodentro.edu.gr/lor/
British Museum and Department for Education, UK: Teaching history with 100 objects http://www.teachinghistory100.org/
CIDOC-CRM http://www.cidoc-crm.org/
EUROPEANA http://www.europeana.eu/portal/
GameIt http://www.gameit.gr/
Getty Vocabularies
http://www.getty.edu/research/tools /vocabularies/
LOM https://standards.ieee.org/findstds/standard/ 1484.12.1-2002.html
LRE http://lreforschools.eun.org/
SPECTRUM http://www.collectionstrust.org.uk/spectrum
Vocabulary Bank for Education – VBE http://aspect.vocman.com/vbe/home

Arts and Cultural Education at School in Europe Education. Audiovisual and Culture Executive Agency (EACEA P9 Eurydice). (2009). http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/113EN.pdf

Fletcher, J. (2014). Gamifying the museum: educational games, museum pedagogy and curriculum. AIS ICT Integration Conference, September 2014, Sydney.

Lambropoulos, N., Christopoulou, M. and Vlachos, K. (2005). Cultural-based Language Learning Objects in CALL. In Zaphiris, P. and Zacharia, G. (2005*) CALL: Computer Assisted Language Learning*. Hershey, PA, USA: Idea Publishing.

Le Boeuf, P., Doerr, M., Ore, C. E., Stead, S. (main editors). (2015). CIDOC Conceptual Reference Manual.

Massart, D. and Shulman, E. (2011). Learning Resource Exchange Metadata Application Profile version 4.7, European Schoolnet, 2011.

Sowa, J.F. (2009). Building, Sharing, and Merging Ontologies. http://www.jfsowa.com/ontology/ontoshar.htm