Documenting a fruitstand from Karabournaki with CIDOC/CRM

Maria Sitara[1], Marios Poulos[2], Sozon Papavlasopoulos[3]

[1]Archaeologist, Postgraduate student at Ionian University, Department of Archives, Library Science & Museum Studies. Ministry of Culture and Sports, Ephorate of Antiquities at Corfu, Old Fortress, 49100 Corfu, Greece, l12sita@ionio.gr.
[2, 3]Assistant Professors at Ionian University, Department of Archives, Library Science & Museum Studies, Ionian University, I. Theotoki 72, 49100 Corfu, Greece, mpoulos@ionio.gr, sozon@ionio.gr

Abstract. The documentation process of cultural goods, originated from Museums, Libraries and Archives, has proved to be a very difficult and complex task due to the enormous mass of data and the variety of the documentation standards that each organization uses. In this paper, we present the design of an ideal model for the documentation of archaeological objects based on CIDOC CRM guidelines for museum documentation, that occurred during our effort to document a fruitstand from a systematic excavation that takes place at Karabournaki, during the last decade.
Keywords: Cultural Heritage, Archaeological Data, Documentation, Ontologies, CIDOC/CRM, Fruitstand K97.1007

1. Introduction

The technological developments of the last decades have promoted new perceptions of the importance and role of culture in an ever-changing society, in which the diffusion of information takes one of the leading positions. The international community has created Institutions for the management of Cultural Heritage and Memory Organizations that have as their primary objective the preservation, conservation, study and interpretation, but also the availability of the collections that they hold (Dempsey, 1999).
Documentation, in other words the process of collecting and recording data, holds a prominent role in the management of cultural goods, as it offers the full history of each object, providing as much information as possible for the future (Moore, 2001: 1). The data are obtained throughout recording cards, photos, notes, fieldwork diaries, correspondence, legal documents, drawings, books,

Received: 21.3.2015 / Accepted: 1.5.2015 ISSN 2241-1925
© ISAST

sound recordings, video recordings, etc. It is obvious, therefore, that this is an ongoing process (Bounia, 2012: 77-78; Kalamara et al., 2004: 2).

Cultural Heritage Organizations have at their disposal various collections of different material, derived not only from Museums, but also from Libraries and Archives, to which they provide access to through the output of metadata. In recent years documentation acquired an international character, while a more systematic effort to record data has been promoted in order to ensure the completeness and uniformity/homogeneity of information, by creating common descriptive standards that provide specific and comparable concepts, terms and procedures in order to optimize the efficiency of documentation, contributing to the smooth exchange of data between various systems (Ghosh et al., 2006; Naumann et al., 2000). The use of a common reference model for the exchange of information can ensure also the semantic interoperability, increasing the efficiency of a search, as it helps to visualize the relationships between the data and also to unify and retrieve the information from different sources. The aim now is not the simple identification of words, but of concepts, while the search results will be based on the meaning of words (Mavromichali, 2006-2007: 17-20). The tool for the representation of concepts and their relationships, but also for the achievement of interoperability is ontologies. "An ontology is a formal, explicit specification of a shared conceptualization of a particular field of interest" (Gruber, 1993: 199). Ontologies constitute a typical representation of a set of concepts expressed in a particular field, as well as the relationships between them. Regardless of the language used for their encoding, most ontologies have some basic structural similarities. Particularly, **Classes**, that are represented by concepts, in their broad meaning, **Relations**, that express a kind of interaction between the concepts of a field, **Functions**, which represent a special case of relation, and **Axioms**, that are used to represent sentences that are always true. Finally, there are **Instances** that express specific data (Mavromichali, 2006-2007: 23). A typical example of a current descriptive standard for museum cultural objects is the ontology CIDOC CRM [ICOM/CIDOC, CRM SIG], to which we shall refer afterwards.

The remainder of this paper is structured as follows. In **Section 2**, we provide a brief description of the Cidoc CRM model, which constitutes the basis of the creation of our model. In **Section 3**, we present the current documentation practice for the archaeological objects. In **Section 4**, we present the ideal documentation model that we created for the recording of movable objects. In **section 5**, we provide a brief description of the case study, followed by the corellation of our case study’s data to the created model and we conclude in **Section 6**, by discussing the pros and cons of the whole procedure, as well as presenting our future work.

2. CIDOC CRM.

One of the main international standards for the classification of the heterogeneous cultural content of museums, libraries and archives is the object-oriented semantic model CIDOC/CRM. This is an event-central ontology based on the axis of time, developed by CIDOC in cooperation with the Institute of

Computer Science of the Foundation of Research & Technology-Hellas (ICOM/CIDOC, 2015; Doerr at al., 2007: 51-56) that provides an ideal analysis of cultural documentation in accordance with the rules of logic, covering every type of material, suggesting a common base for interpretation without imposing any directions as to what should be documented, while at the same time it can be used as an information exchange protocol. It is intended to facilitate the integration, mediation and interchange of heterogeneous cultural heritage information (Constantopoulos et al., 2005: 3; Lourdi, 2010: 67-68; Mavromichali 2006-2007: 33; Crofts et al., 2003). CIDOC CRM represents a higher level of abstraction, categorizing the cultural heritage data under general groups, such as Persons, Places, Time, Activities, and conceptualizing the relationship established between these groups. It includes approximately 89 classes and 151 properties and covers the semantic information of hundreds of schemata. Classes form a predetermined hierarchy from the general to the specific term. It is designed to be extensible through the linkage of compatible external type hierarchies, making it potentially infinite (ICOM/CIDOC, 2015).

3. Current Documentation Practice for the Archaeological Objects.

Up to now documentation was realized either in manuscript form or electronically, based on data bases, whose main goal was to organize all the information so that they could be easily accessed, managed and updated. The problem that occurred was that each organization used different data bases according to its management policy, thus, leading to data redundancy and inconsistency, as well as to an inability regarding data sharing and standardization (Bounia, 2012: 77-78; Elings & Waibel, 2007; Lourdi, 2010: 26). Several attempts were made in order to solve the problems of heterogeneous data schemes and processing in archaeology, one of which was the ETANA-DL unified (meta)-model for archaeological systems based on the 5S framework for information systems, using the client-server paradigm of the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). It supported services such as: searching, browsing, personal collections, workflows, discussions, annotations, personal lists, recording of most recent queries issued, etc., that could be reused in other DLs, and also could be combined to create a full-fledged DL without compromising on performance (Ravindranathan et al., 2004: 76-77).

However, besides the differences between the various databases which are used currently by the organizations, the structure for the documentation of objects has to include some basic common elements, such as:

- The **general data** concerning the object. That is the object's identifier, the actor and date of the documentation, the type of the object and its description, the technical characteristics, the material and its condition assessment.

- **Dimensions**, which can be analyzed to: length, height, width, base/ rim/ handle diameter, maximum diameter, base/ rim thickness, maximum/ minimum thickness.
- The **spatial – temporal data**, such as: the production workshop, potter/ painter, the dating, the date of the discovery, the excavational section/ area, the geospatial coordinates and the depth of the discovery.
- The data regarding the **description** of the object, such as: the full description, the inscriptions, the similarity with other objects and generally comments/ remarks concerning data that cannot be included in any other field.
- And finally, the **bibliography** that was used in order to identify the type and possibly the origin of the object and its dating, consisting of data such as: author(s), publication, book/ journal, pages.

All these data can be supplemented with photos, slides and designs. Also a reference can be made to the objects that were found along with the object that is being described. It should be noted that information about the modifications, the move or the exhibition of an object are documented in different periods and data bases.

4. Design of a Documentation Model for non-movable objects.

An ideal documentation model for excavation findings was created, based on the guidelines of the CIDOC model for Museum Documentation, which can be supplemented with more entities or data. The structure of our model was divided into two main categories: the events of the "ancient" history and the events of the "modern" history of an object. The main questions that are raised in most events are related to issues concerning location (where), time (when), actors (who), causes (why) and classification (types). As a result, a top-down class hierarchy is formed, after defining the general concepts, followed by their specialization. After the definition of several classes, the description of the internal structure of the concepts begins, in order to identify the relationships between the terms. These are the properties of the classes. Then follows the definition of the characteristics of these properties that are associated with the type of values that they can receive (full text, string, numbers, integers, etc.), the cardinality of values etc. In more details, an example of the Entity Object (E19, E22) is given as follows:

The Entity Object (E19, E22) includes the basic information needed to identify, define and classify an object inside a set of similar objects, such as: a unique identifier ID (P47→E42, 0,n:0,n), title (P102→E35, 0,n:0,n), condition state (P44→E3, 0,n:1,1) and number of parts (P57→E60, 0,1:0,n), section definition (P58→E46, 0,n:1,1), type (if it is simple/plain, complex, or set) (P2→E55, 0,n:0,n), category (P2→E55, 0,n:0,n), type of use (e.g. vase etc., P2→E55, 0,n:0,n), type of material (in terms of material, e.g., clay P45→E57, 1,n:0,n), type of colour (P2→E55, 0,n:0,n), stylistic type (e.g. skyfoeides P2→E55, 0,n:0,n), and functional type (P103→E55, 0,n:0,n). For each object there is also the possibility of developing an informational text (P3→E62, 0,n:0,1). In a same way the other entities are connected such as these are depicted in Figure 1.

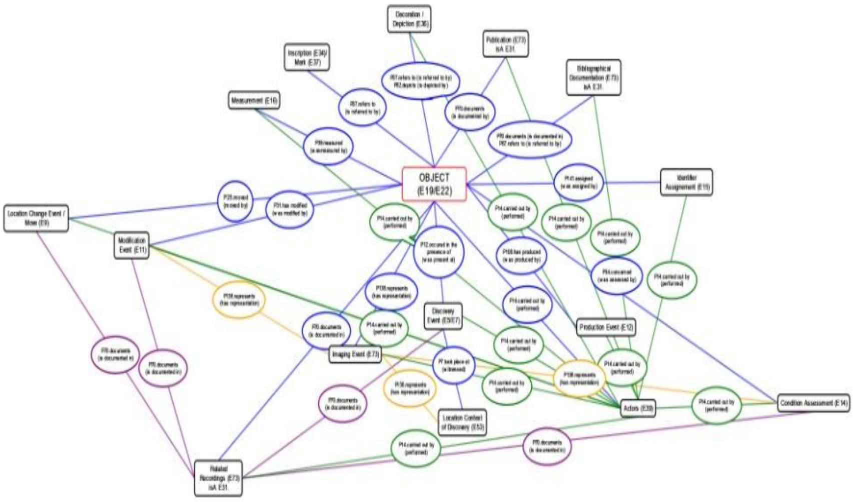

Figure 1. CRM model.

The final phase regarding the design of our ontology is the entry of the classes' instances that represent specific data, existing objects, as well as abstract entities of the field of interest that our ontology is covering.

5. Case-Study: Fruitstand No K97.1007.

The case-study of this paper is the fruitstand No K97.1007, which comes from the on-going excavations in Karabournaki Toumba that is being carried out from 1994 up to today from the Aristotle University of Thessaloniki under the supervision of Professor M. Tiberius and his collaborators, Mrs. E. Manakidou and Mrs. D. Tsiafaki. This eastern Greece fruitstand, from a Milesian workshop, was detected in 1997 in three fragments that preserve a quarter of the main body of the vessel. It is covered with a white coating and is decorated with geometric motifs, of black, violet and white colourandis dated between 610 - 580/570 (Tsiafakis, 2012: 153-154, 156).

In this paper the documentation of the fruitstand is based on two recordings, one from the Excavation Data Base and the other from the Web Data Base, a 3D reconstruction model, and three publications concerning the mentioned object. We noticed that in each of the aforementioned sources some data were repeated, whilst other were unique. As a result, the scientific documentation of the fruitstand seemed incomplete. Thus, we tried to homogenise all our sources in order to ensure the completeness of the documentation.

The two data bases (→E73 isA E31→P70→E1/E7/E9/E11/E14 E19/E22 (1,n:0,n)→title→P102→E35 (0,n:0,n)→recording category→P2→E55 (0,n:0,n)→time span→P4→E52 (1,1:1,n)→recording identifier→P1→E41

(0,n:0,n)→language→P72→E56 (0,n:0,n)→actors→P14→E39, P14.1→E55 (1,n:0,n) used to describe Fruitstand K97.1007 follow the current documentation structure, described earlier. A full recording of all the information coming from the heterogeneous "sources" was realized. Afterwards, all the information was organized and discerned into the individual sections of the base, and were correlated with our CIDOC model, as it is presented below:

- G**eneral data**: object's identifier→unique identifier ID)→E19/E22→ P47→E42→has type→P2→E55 (0,n:0,n), the actor→actors→P14→ E39→in the role of→P14.1→E55 (1,n:0,n) and date of the documentation) →time span→P4→E52 (1,1:1,n), the type of the object that is being documented→stylistic type→P2→E55 (0,n:0,n), and its description→title→ P102→E35 (0,n:0,n), the technical characteristics, the material→type of material→P45→E57 (1,n:0,n) and type of colour→P2→E55 (0,n:0,n), and its condition assessment→condition state→P44→E3 (0,n:1,1), has type→ P2→E55 (0,n:0,n).
- **Dimensions**→Measurement→E16→P39→E1 (1,1:0,n) which can be analyzed to: length, height, width, base/ rim/ handle diameter, maximum diameter, base/ rim thickness, maximum/ minimum thickness→ type of measurement→P40→E54 (1,n:0,n), measurement unit→P91→E58 (1,1:0,n), value→P90→E60 (1,1:0,n).
- The **spatial – temporal data**, such as: the production workshop→Event of Production→E12→P108→E19/E22 (1,n:1,1)→place→ P7→E53 (1,n:0,n), potter/ painter→actors→P14→E39→in the role of→P14.1→E55 (1,n:0,n), the dating→time of production→P4→E52 (1,1:1,n), the date of the discovery→Discovery Event→E5/E7→P12→E77/E19/E22 (1,n:0,n)→time span of the discovery→P4→E52 (1,1:1,n), the excavational section/ area→Location Context of the Discovery→E53→P7→E7 (1,n:0,n) →place appellation→P87→E44, P89→E53 (0,n:0,n)→type→P2→ E55 (0,n:0,n), the geospatial coordinates and the depth of the discovery→spatial coordinates→P87→E47 (0,n:0,n)→layer of the objects discovery→ P87→E44/E42 (0,n:0,n).
- The data regarding the **description** of the object, such as: the full description→Decoration/Depiction→E36→P67→E1, or E36→P62→E1 (0,n:0,n)→content→P138→E1 (0,n:0,n), the inscriptions→Inscription/ Marks→ E34/E37→P67→E1 (0,n:0,n), the similarity with other objects→ Bibliographical Documentation→E73 isA E31→P70→E1/E19/E22 (1,n:0,n) or E73 isA E31→P67→E1/E19/E22 (0,n:0,n)→title→P102→E35 (0,n:0,n) →bibliographical references→P67→ E1 (0,n:0,n)→ objects in comparison→ E70→P130→E70, P130.1→E55 (0,n:0,n)and generally comments/ remarks concerning data that cannot be included in any other field→informational text→P3→E62 (0,n:0,1).
- And finally, the **bibliography** →Bibliographical Documentation→E73 isA E31→P70→E1/E19/E22 (1,n:0,n) or E73 isA

E31→P67→E1/E19/E22 (0,n:0,n)→title→P102→E35 (0,n:0,n), that was used in order to identify the type and possibly the origin of the object and its dating, consisting data such as: publication→time span→P4→E52 (1,1:1,n)→place→P7→E53 (1,n:0,n) →publication identifier→P1→E41 (0,n:0,n), book/journal→bibliographical category→P2→E55 (0,n:0,n), pages→bibliographical references→P67→E1 (0,n:0,n) and actor→actors→P14→ E39→in the role of→P14.1→E55 (1,n:0,n).

All these data were supplemented with photos, slides, designs and a 3D reconstruction model→Imaging Event→E73→P138→E1/E7/E11/E14/E19/E36/ E53, P138.1→E55 (0,n:0,n)→visual item used→P2→E55 (0,n:0,n)→type of imaging→P2→E55 (0,n:0,n)→appellation→P102→E35 (0,n:0,n).

6. Conclusions.

The creation process of the aforementioned model proved to be arduous, due to the huge, and therefore hardly manageable, mass of information that needed to be sorted out and classified to categories. Some data were repeated, others overlapped each other, and there were cases where different terms were used to elaborate the same or similar fields. It is obvious that such documentation must be conducted by experts, as it needs necessarily knowledge of the archaeological terminology and also of the excavation and documentation process.

The presented documentation model can be supplemented with new entities regarding new or other events of the object's life, additional data (instances) to the existing entities, whilst offer the possibility to amplify the already existing entities, according to our field of interest. This documentation model is useful to researchers because:

- It provides all the data concerning an object, along with information regarding the spatial and temporal context in which the finding is placed.
- It presents the relationship with the related objects found within the same non-movable structure and the same layer, and thus can lead to conclusions about the use and dating of the site.
- Through the bibliographical references it provides the similarity with other objects and the connection with the production workshops, presenting in this way the dispersion in the geographical area.
- It also offers all the data concerning the actors involved in the various 'life' events of the object.
- And last but not least, it is able to provide data regarding to the effect of the object's discovery, which can lead to publications, exhibitions etc.

Future work includes the addition of further data relating to the modification and the exhibition events of the object, as well as all the data concerning the related objects that were found in the same area with the fruitstand. Furthermore, the implementation of the CRM documentation model to the open source platform Protégé will follow, in order to create a full mapping of our ontology (Protégé).

References

Bounia, A. (2012), Behind the scenes of the museum: The management of museum collections, edt. Patakis, Athens.

Constantopoulos, P., Bekiari, Ch., Doerr, M. (2005), Guide for the Cultural Documentation and Interoperability. Interoperability and Access, (in Greek), Foundation of Research & Technology-Hellas, Institute of Computer Science, Athens.

Crofts, N., Doerr, M., & Gill, T. (2003), The CIDOC conceptual reference model: A standard for communicating cultural contents, in Cultivate Interactive, 9.

Dempsey, L. (1999), Scientific, industrial, and cultural heritage: A shared approach: A research framework for digital libraries, museums and archives, in Ariadne, 22.

Doerr, M., Ore Ch.-E., Stead, St. (2007), The CIDOC Conceptual Reference Model – A New Standard for Knoeledge Sharing, in Tutorials, Posters, Panels and Industrial Contributions at the 26th International Conference on Conceptual Modeling-ER 2007; CRPIT. Auckland, New Zealand.

Elings, M. W. & Waibel, G. (2007), Metadata for All: Descriptive Standards and Metadata Sharing across Libraries, Archives and Museums, in First Monday, 12(3).

Ghosh, S.K. & Manoj, P. (2006), Geo-Spatial Interoperability: Crossing Semantic and Syntactic barrier in GIS, in GIS Asia Pacific, 10(9).

Gruber T.R (1993), Toward principles for the design of ontologies used for knowledge sharing, in International Journal Human-Computer Studies, 43.

ICOM/CIDOC, Documentation Standards Working Group & CIDOC CRM SIG. Available in<http://network.icom.museum/cidoc/working-groups/crm-special-interest-group/> (Last accessed 10.03.2015)

ICOM/CIDOC, Documentation Standards Group & CIDOC CRM SIG (2015), Definition of the CIDOC Conceptual Reference Model: New version of the CIDOC Conceptual Reference Model: Version 6.0 of the reference document (January 2015). Available in <http://cidoc.ics.forth.gr/official_release_cidoc.html> (Last accessed 10.03.2015)

Kalamara, P, Vagelatos, A. (2004), 'Twists' concening the cultural heritage management: The dynamic and boundaries of new technologies, in 2nd International Conference on Museology, (in Greek), Mytilene.

Lourdi, E. (2010), Designing a metadata model to document digitized Folklore Collections, (Doctoral Dissertation), (in Greek), Ionian University, Department of Archives and Library Science, Corfu.

Mavromichali, M. (2006-2007), The Description of museum collections, (Master Thesis), (in Greek), Ionian University, Department of Archives and Library Science, Corfu.

Moore, M., (2001), Conservation Digitization and the Implications of Digitisation, in Journal of Conservation and Museum Studies, (7).

Naumann, F., Freytag, J., & Leser, U. (2000), Completeness of Information Sources, in Proceedings of Ulf Leser Workshop on Data Quality in Cooperative Information Systems (DQCIS) 2003, Heidelberg: Springer-Verlag (ed.).

Protégé, available in <http://protege.stanford.edu/> (Last accessed 10.03.2015)

Ravindranathan, U., Shen, R., Gonçalves, M. A., Fan, W., Fox, E. A., Flanagan, J. W. (2004), ETANA-DL: A Digital Library For Integrated Handling Of Heterogeneous Archaeological Data, in Proceedings of the 2004 Joint ACM/IEEE Conference on Digital Libraries (JCDL'04), New York.

Tsiafakis, D. (2012), An East Greek Fruitstand at Karabournaki, in Kerameos Paides: Studies offered to Professor Michalis Tiverios by his students, (in Greek) (edts) Eurydice Kefalidou & Despoina Tsiafakis, Thessaloniki.