SEMANTIC INTEGRATION OF COLLECTION-LEVEL INFORMATION: A CROSSWALK BETWEEN CIDOC/CRM & DUBLIN CORE COLLECTIONS APPLICATION PROFILE

Irene Lourdi
Libraries Computer Centre
National & Kapodistrian University of Athens
University Campus, Ilisia, Dept. of Informatics & Telecommunications
Athens
Greece
E-Mail: elourdi@lib.uoa.gr

Christos Papatheodorou
Dept. of Archives and Library Sciences and Digital Curation Unit
Ionian University and "Athena" Research Centre
72 Ioannou Theotoki str., 49100
Corfu
Greece
E-Mail: papatheodor@ionio.gr
URL: http://www.ionio.gr/~papatheodor

Abstract

This paper is motivated by the demand for unified access, navigation and information retrieval from the wealth of composite, distributed and heterogeneous digital cultural collections. The last years, collection-level metadata is considered to be the key of integrated access of so many resources, since they represent the inherent and contextual characteristics of a collection. Our effort origins from the semantic interoperability perspective and considers CIDOC/CRM as the mediating schema, which integrates in an optimal way the semantics of the collection level metadata schemas and application profiles. In particular a crosswalk between Dublin Core Collections Application Profile and CIDOC/CRM is presented so that the semantics of each DCCAP element is mapped to CIDOC/CRM. The derived crosswalk is bidirectional implementing the mapping from DCCAP to CIDOC/CRM and vice versa. The paper reveals the complexity of mapping metadata schemas to ontologies and resolves particular difficulties providing a real world semantic integration case.

INTRODUCTION

The last years an explosion of cultural heritage digitization projects has been implemented taken place with the result to prevail a plethora of information in the web, which it is considered the central forum for data storage and information exchange between various digital applications. According to Gill and Miller (2002), digital cultural content is a broad concept that includes multimedia surrogates for material owned by the world's museums, libraries and archives, and their associated descriptive and contextual information. The implementation of all these digitization projects aims not only at material preservation but also at getting people close to cultural heritage.

The information landscape has changed recently. Researchers and simple users desire to find intellectual and cultural materials easily, without concerning about institutional or national boundaries (Dempsey, 2000). When users need to find information for a specific subject, they prefer to collect data from all the institutions and sources, which might be holders of this kind of information, through an integrated digital environment. This kind of request is quite difficult to be achieved, since all the digitization projects use various metadata schemas according their particular resource description needs. The existing metadata heterogeneity makes difficult for the user to access the information he is seeking, and quite often he is confused or miss valuable information for his own interest.

Thus, the so–called "memory institutions" (Dempsey, 2000) in front of the increasing demand for global access to highly distributed, heterogeneous, and dynamic cultural heritage wealth are exploring possible solutions for integrating and analyzing the data from multiple sources. They are obliged to provide their cultural resources in a unified manner, to ensure that users will follow unified routes to their goals. In this context scientific research has focused on matters of interoperability and integration, taking into account that many cultural heritage metadata schemas, even though they differ in format, language or structure, they are often semantically related with each other.

To face this challenge, in this paper we propose an efficient ontology-based method to integrate the collection-level metadata, giving to the users the chance to find easily information not only about item-level descriptions but also about the collections. In particular we present a crosswalk between Dublin Core Collections Application Profile (DCCAP) (Dublin, 2006) and the CIDOC/CRM (Icom/Cidoc, 2006) by focusing on important mapping issues. The harmonization of these two schemas is expected to offer a semantically rich model able to present historical events and facilitate the conceptual unification of the cultural collections' entities as well as the information retrieval from various sources.

The outline of the paper is as follows. The following section provides an overview of the collection-level metadata schemas and the related work on information integration. Then the cultural heritage collection curation activities are represented using CIDOC/CRM, while in the fourth section the DCCAP model is briefly presented. In the next section the mapping between DCCAP and CIDOC/CRM is presented, addressing the main issues the mapping confronts and finally the paper closes with the presentation of the conclusions of the whole research effort.

SEMANTIC INTEGRATION OF COLLECTION–LEVEL METADATA

Collection – level metadata schemas

The need for collection-level descriptions became urgent only when there have been a huge volume of remote databases and different digital applications. Except archives, libraries and museums focused their cataloguing on item-level and therefore until recently there have been no best practices and cataloguing rules for the description of collection-level entities. The realm of digital libraries paid particular attention on the collection-level description due to the volume of available digital content and need for the integration of the distributed and heterogeneous information.

Cultural heritage institutions have developed various metadata schemas for collection-level description satisfying their institution management policy and requirements. The diversity of institutional management perspectives, makes impossible to create a single descriptive schema covering almost all the communities needs (Gill, 2004).

Hence a plethora of standards and local community-specific metadata models have been created employing different data structures, data content rules and (to some extent) data formats to encode their collections and the information they contain. The most widely known standards for collection-level description are: RSLP (Powell), DCCAP (Dublin, 2006), NISO (ANSI/NISO) and EAD (LC, 2002). All these schemas are targeted on specific types of organizations, for example EAD on archives and RSLP on library collections, but they have the same scope of description.

In general, collection-level description schemas are not a substitute for item-level description (Chapman, 2004). They really intend to allow users to search for information across library, museum and archive domains and they are clearly desirable, since they enable the discovery of collections of interest, prior to item-level discovery. Collection-level metadata represent the inherent and contextual characteristics of a collection and are considered a key factor for gathering information from distributed digital resources. As Hakala in (2005) claims, "collection-level descriptions facilitate an agent to:

- Discover collections of potential interest
- Identify a collection
- Select one or more collections from amongst a number of discovered collections
- Identify the informational services that provide access to the collection"

Interoperability and Semantic integration approaches

From the very start, the ability to search across a range of various resources and bring together collections was a high priority for digital libraries. A number of authors have attempted to define *interoperability*. Hunter (2001) claims that interoperability is intended to "enable a single search interface across heterogeneous metadata descriptions, to enable the integration or merging of descriptions which are based on complementary but possibly overlapping metadata schemas or standards and to enable different views of the one underlying and complete metadata description, depending on the user's particular interest, perspective or requirements".

A widely used method for metadata interoperability is *crosswalks* between various metadata schemas. Many papers have been written for metadata mappings and crosswalks. A very significant paper, that delineates the general issues involved in the harmonization of metadata standards and in the development of crosswalks between related metadata standards, is of the National Information Standards Organization (NISO) called "Issues in crosswalking content metadata standards" (Pierre, 1998). Further, a good discussion of crosswalking issues can be found in the paper of Mary Woodley (2000). Another good reference guide for metadata mappings is available by the MIT Libraries (2007) and the UK Office for Library and Information Networking (Day 1996).

Recently research interests are moving to *semantic integration* of heterogeneous resources, focusing especially on the cultural heritage domain, for which there exist many and complex metadata schemas. According to Cruz and Xiao (2005), "semantic integration defines and uses conceptual representations of the data and their relationships in order to eliminate possible heterogeneities", meaning that in the diverse digital environment, semantic integration reveals and makes machine understandable the rich semantics of the resources, facilitating information integration. Efforts, that have been made to develop semantic approaches for data integration and are valuable to be referred, are: the project of Knowledge Sharing Effort (Darpa, 1990) and Knowledge-based Integration (Nam et al. 2002). In (Tous, 2006), a model-mapping approach is applied to represent instances of XML and XML Schema in RDF. The described architecture proposes a semantic XPath processor that acts over an RDF mapping of XML and is fed with an unlimited set of XML schemas and/or RDFS/OWL ontologies.

In many of these approaches the mapping effectiveness has not yet been examined for really complex data structures like metadata schemas. In (Amann et al. 2001) a mechanism for the cultural information sources integration is proposed. The authors map pieces of information contained in XML fragments to domain specific ontologies, such as CIDOC, defining (1) a mapping language that describes the resources by a set of rules relating XPath location paths to the concepts and roles of an ontology, and (2) a query rewriting algorithm for translating user queries into queries expressed in an XML query language, which are send for evaluation to XML sources.

According to (Hong et al. 2005), "physically combining data into a single system may be impossible for technical, organizational or economic reasons". Thus he proposes the utilization of a typical mediation system, which will act as a single interface for users. A mediation schema could be consider the CIDOC/CRM, which in (Crofts 2003) acts as the global schema allowing the gathering of all the necessary cultural information in a suitable form for further reasoning. Nevertheless, there are quite many works referring in mapping data schemas to CIDOC/CRM that try to enable the exchange and sharing of heterogeneous sources both within and between cultural institutions. Most of them can be found in the official site of CIDOC/CRM[1]. For instance, Hunter (2002) presents an effort that combines MPEG-7 with CIDOC/CRM into a single ontology aiming to describe and manipulate cultural multimedia resources. In this case, CIDOC/CRM is extended by MPEG-7 components, to gain multimedia metadata manipulation capabilities. Further, Doerr (2000) presents a mapping of the Dublin Core Metadata Element Set to the CIDOC/CRM, while this paper extends and refines the mapping presented in (Kakali et. al. 2007) focusing on the crosswalk between DCCAP and CIDOC/CRM.

The rationale of the proposed crosswalk

According to Doerr and LeBoeuf (2007) "*Core ontologies describing the semantics of metadata schemata are the most effective tool to drive global schema and information integration, and provide a more robust, scalable solution than tailored 'cross-walks' between individual schemata*". Hence, instead of mapping various metadata schemas with each other, the authors propose to map them to core ontology.

In our approach the integration of collection-level metadata requires the development of crosswalks from the various schemas to CIDOC/CRM ontology. In this context

[1] CIDOC CRM Special Interest Group, Working Group of CIDOC. CIDOC CRM Mappings, Specializations and Data Examples/ http://cidoc.ics.forth.gr/crm_mappings.html

CIDOC/CRM has been selected as the mediating schema because it defines the complex interrelationships that exist between objects, actors, events, places and other concepts in the cultural heritage field. The value of CIDOC/CRM becomes apparent when it is used as the basis for data transfer and exchange between different systems, schemas and semantics (Crofts 2003). CIDOC/CRM could serve as a virtual global schema and has the capability to integrate complementary information from more restricted schemata (Doerr, 2006). On the other side DCCAP was chosen as the first collection-level schema to be mapped to CIDOC/CRM, since it is based on a widely accepted metadata set that facilitates information discovery and is applied for interoperability reasons by many communities. The challenge to facilitate a sufficiently expressive model for extracting information from collection-level descriptions led us to the effort of harmonizing conceptualizations from relevant domains.

CIDOC/CRM: REPRESENTING A CULTURAL HERITAGE COLLECTION

The CIDOC Conceptual Reference Model (CRM) has been developed since 1996 under the auspices of the International Committee on Documentation (CIDOC) of the International Council for Museums (ICOM) Documentation Standards Working Group. It is a core ontology aiming to integrate the concepts of cultural heritage documentation and facilitate the controlled exchange of the information between various memory institutions like archives, libraries and museums. Its main characteristic is that the descriptions do not focus on the objects but on the events that connect people, material and immaterial things in space-time (Doerr, 2006). The usage of CIDOC/CRM ontology enhances accessibility to museum-related information and knowledge and provides an important information standard and reference model for Semantic Web initiatives.

CIDOC/CRM ontology consists of 81 classes and 132 unique properties. Each class is identified by numbers preceded by the letter “E” (classes are referred to as “Entities”), and are named using noun phrases (nominal groups) using title case (initial capitals). For example, the class E39 is named “Actor”. Each property has a domain and range class is considered symmetric and is identified by numbers preceded by the letter "P" and named using verbal phrases in lower case. In general terms, the CIDOC/CRM classes and property hierarchies have the ability to describe various kind of information, like: a) Identification, Acquisition and ownership information, b) Physical movement, location and relocation information, c) Physical attributes and features and d) Historical events: both real world events, as well as the events which occur in the life cycle of a resource, and those events which are depicted in the visual information objects.

General characteristics	
Collection Name	E78(Collection)- *P102 has title (is title of)*- E35(Title)
Collection Identifier	E78(Collection)- *P47 is identified by*- E42(Object identifier)
Collection Size	E78(Collection)- *P57 has number of parts*- E60(Number)
Location	E78(Collection)- *P53 has former or current location (is former or current location of)*-E53(Place) E78(Collection)- *P25 moved (moved by)*- E9(Move)- *P26 moved to (was destination of)*- E53(Place)
Rights held over collection	E78(Collection)- *P104 is subject to*- E30(Right)
General notes	E78(Collection)- *P3 has note*- E62(String)
Description	E78(Collection)- *P62 depicts (is depicted by)*- E1(CRM Entity)
Associated object	E78(Collection)-*P130 shows features of*-E70(Thing)
Associated publication	E78(Collection)-*P67 refers to (is referred to by)*- E73(Information Object)
Persons related with the collection	
In CIDOC/CRM ontology persons are presented with the class E39 Person. This class comprises people, either individually or in groups, who perform intentional actions for which they can be held responsible. Therefore persons can only be described through the Activity to which they participate.	
Curator	E78(Collection)- *P109 has current or former curator (is current or former curator of)*- E39(Actor)
Owner	E78(Collection)- *P52(has current owner/is current owner)*- E39(Actor)
Keeper	E78(Collection)- *P49 has former or current keeper (is former or current keeper of)*- E39(Actor)
Person/body responsible for transfers	E78(Collection)- *P25 moved (moved by)*- E9(Move)- *P14 carried out by (performed)*- E39(Actor)
Person/body responsible for destructions	E78(Collection)- *P13 destroyed (was destroyed by)*- E6(Destruction)- *P14 carried out by (performed)*- E39(Actor)
The Events that have affected the collection	
Curation Activity	E78(Collection)- *P147 curated (was curated by)*- E87(Curation Activity)
Acquisition Event	E78(Collection)- *P24 transferred title of (change ownership through)*-E8(Acquisition)
Move	E78(Collection)- *P25 moved (moved by)*- E9(Move)
Transfer of Custody	E78(Collection)- *P30 transferred custody of (custody transferred*

	through)-E10(Transfer of Custody)
End of Existence	E78(Collection)- *P93 took out of existence (was taken out of existence by)*-E64(End of Existence)
Items or parts of the collections A collection, as it is said above, contains aggregations of physical items. Usually a collection-level description does not intend to cover the description of containing objects, but gives an integrated picture of the collection. If it is necessary to give more details about the collection components the corresponding path is: E78(Collection)-*P46 is composed of (forms part of)*-E18 Physical Thing	

Table 1. CIDOC entities defined for E78 Collection

The term "cultural heritage collections" refers to any materials collected by museums, as defined in the ICOM statutes (ICOM 1946–2001). The collection entity in CIDOC/CRM is presented by the class "E78 Collection", which is a subclass of "E24 Physical Man-Made Thing" and according to the definition "*it comprises aggregations of physical items that are assembled and maintained ("curated" and "preserved," in museological terminology) by one or more instances of "E39 Actor" over time for a specific purpose and audience, and according to a particular collection development plan*". A collection can be treated and described like a separate cultural object, for which CIDOC/CRM covers the historical, geographical and theoretical context information.

To model cultural heritage collections using CIDOC/CRM, we firstly create all the possible CIDOC/CRM paths in which the "E78 Collection" class participates either as domain or range. A CIDOC/CRM path has been clearly defined according to (Kondylakis et al. 2006) as a triple of the form *class-property-class* (*c-p-c*), such that the clases of the triple correspond to the property's domain and range. For example, consider the CIDOC/CRM path E78(Collection) - *P47 is identified* - E42(Object identifier). This path denotes the unique identifier of the described Collection.

In general terms, a collection is characterized by the following issues: a) the persons related with the collection, b) the collection's main characteristics (name, identity, location etc.), c) the objects that the collection contains and d) the various events that affected the collection (update, reallocation, etc.). The CIDOC/CRM paths, that express this kind of information, are presented in Table 1.

DUBLIN CORE COLLECTIONS APPLICATION PROFILE DESCRIPTION

Dublin Core Collections Application Profile (DCCAP) is an application profile for collection-level description. As it is described by the Dublin Core Collection Description Task Group, "*it provides a means of creating simple descriptions of collections suitable for a broad range of collections, as well as simple descriptions of catalogues and indexes*". It is considered to be valuable for extracting a set of core data from a resource and for providing users important information about a collection. A collection can be "*any aggregation of physical or digital items, while a collection description includes a description of one or more collections (aggregations of items) and a description of zero or more catalogues or indices (resources which describe collections)*".

The properties of the application profile are presented analytically in Table 2, remarking that DCCAP is still under development. The application profile is based on Dublin Core (Dublin, 2008) as well as various metadata vocabularies. For this reason the elements in Table 2 are specified by the following namespaces:
The Dublin Core Metadata Element Set, v1.1, with the prefix **dc**
Dublin Core Terms, with the prefix **dcterms**
Dublin Core Type Vocabulary, with the prefix **dcmitype**
MARC Relator Code Properties, with the prefix **marcrel**
Collection Description Terms, with the prefix **cld**
Collection Description Type Vocabulary Terms, with the prefix **cdtype**

Property name	
Collection Identifier (dc:identifier)	Accrual Method (dcterms:accrualMethod)
Title (dc:title)	Sub-Collection (dcterms:hasPart)
Alternative Title (dcterms:alternative)	Super-Collection (dcterms:isPartOf)
Description (dcterms:abstract)	Custodial History (dcterms:provenance)
Type (dc:type)	Associated Collection (cld:associatedCollection)
Size (dcterms:extent)	Associated Publication (dcterms:isReferencedBy)

Language (dc:language)	Spatial Coverage (dcterms:spatial)
Item Type (cld:itemType)	Temporal Coverage (dcterms:temporal)
Item Format (cld:itemFormat)	Subject (dc:subject)
Collector (dc:creator)	Dates Collection Accumulated (dcterms:created)
Owner (marcrel:OWN)	Catalogue or Index (cld:catalogueOrIndex)
Rights (dc:rights)	Date Items Created (cld:dateItemsCreated)
Access Rights (dcterms:accessRights)	Is Located At (cld:isLocatedAt)
Audience (dcterms:audience)	Is Accessed Via (cld:isAccessedVia)
Accrual Periodicity (dcterms:accrualPeriodicity)	Accrual Policy (dcterms:accrualPolicy)

Table 2. DCCAP elements

THE CROSSWALK BETWEEN CIDOC/CRM AND DCCAP

Methodology

The process of mapping two or more metadata schemas implies a "table that maps the relationships and equivalencies of them", focusing mainly on a simple semantic correspondence between the elements (Dublin 2001). The creation of mapping is an intellectual work done by domain experts and includes the mapping of each element in the source metadata standard to a semantically equivalent element in the target metadata standard (Pierre 1998).

Contrary to the traditional way of mapping two metadata schemas and since in our paper we deal with the mapping of a metadata set to an ontology model, a different methodology is proposed. According to Kondylakis et al. (2006) "*the mapping of two schemas is as a sufficient specification to transformation of each instance of schema 1 into an instance of schema 2 with the same meaning*". A path-oriented methodology is followed for the development of a crosswalk between DCCAP and CIDOC/CRM, meaning that the ontology paths are mapped to the equivalent metadata paths and vice versa. We remind that we define a CIDOC/CRM path as a triple of the form *class-property-class* (*c-p-c*), such that the classes of the triple correspond to the property's domain and range. Accordingly a DCCAP path is defined as follows: the DCCAP record is linked with a sequence of DCCAP properties. For example, the path DCCAP->cld:IsAccessedVia denotes that there is an electronic service for accessing the collection. Hence, we have to map all the DCCAP paths to semantic equivalent CIDOC/CRM paths and vice versa.

One of the main issues of the proposed crosswalk is the strictly event-aware character of CIDOC/CRM. In particular some classes cannot be associated directly to other classes through a property, but can associated only via the mediation of a sequence of one or more event or activity classes. For instance in Table 1 it is shown that the person/body responsible for the destructions of the objects of a collection (which is mapped to the class E39-Actor), should be associated with the collection class (E78-Collection) through the path E78(Collection)- *P13 destroyed (was destroyed by)*-E6(Destruction)- *P14 carried out by (performed)*- E39(Actor). Thus the crosswalk should express the DCCAP paths to complex CIDOC/CRM paths consisting of sequences of intermediate activities and/or events. For example the path DCCAP-> Accumulation Date range corresponds to the CIDOC/CRM path E78(Collection)-*P108*(*has produced/was produced by)*-E12(Production Event)-*P4(has time span/is time span of)*-E52(Time Span).

Snapshots of the crosswalk

The creation of a crosswalk requires the definition of: a) the mapping between the Source Domain and the Target Domain, b) the mapping between the Source Range and the Target Range, c) the proper Source Path, d) the proper Target Path, e) the mapping between Source Path and Target Path and f) in some cases, the combination of paths sharing the same instances (Kondylakis et al., 2006).

In the following figures 1 and 3 we present two examples of mapping DCCAP paths to CIDOC/CRM paths, demonstrating the main philosophy of the crosswalk. As source domain is considered the DC.Type “Collection” and as a target domain the CIDOC/CRM entity “E78 Collection”.

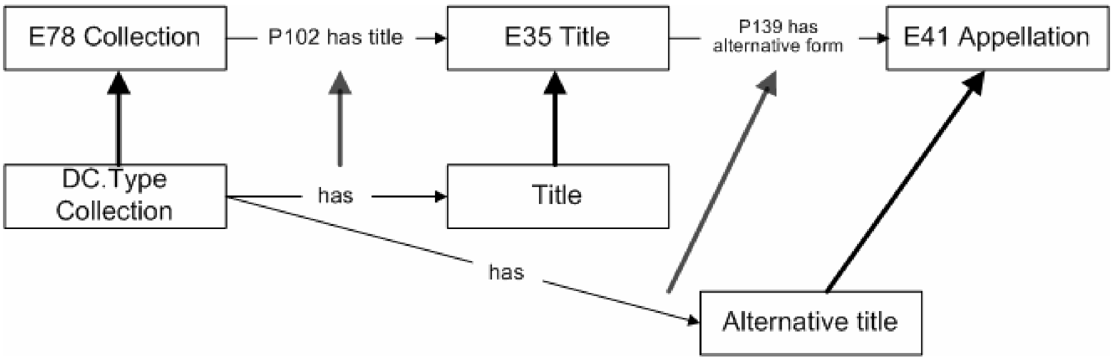

Figure1. Mapping the collection title to CIDOC/CRM

In figure 1 we present the mapping of two DCCAP paths to CIDOC/CRM . Firstly the path DCCAP->dc:title, that corresponds to the CIDOC/CRM path E78(Collection)-*P102(has title)*-E35(Title) and also the path DCCAP->dcterms:alternative that corresponds to the same path but expanded by adding the path E35(Title)-*P139(has alternative form)*-E41(Appelation).

Figure 2 presents an XML implementation of the above mentioned mapping of the path DCCAP->dc:title to a CIDOC/CRM path, based on the XML DTD proposed in (Kondylakis et al., 2006). The DTD presupposes the definition of the source and target domains, which are correspondingly the DCCAP for the source and E78 (Collection) for the target. Each *<link_map>* element describes the mapping of a pair of paths, which belong to different schemas. In particular in this element the source and target ranges are defined which are dc:title and E35 (Title) correspondingly, while the element <path_map> defines the mapping between the source and target paths.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE mapping SYSTEM "C:\Documents and Settings\eirini\Desktop\crm_mapping.dtd">
<mapping>
<map>
<domain_map>
<src_domain>DCCAP</src_domain>
<src_domain_condition>If DC.Type = DCT collection</src_domain_condition>
<target_domain>E78 Collection</target_domain>
<target_domain_condition/>
</domain_map>
<combined_links joined_on="x2">
<link_map>
<range_map>
<src_range>DC.Title</src_range>
<src_range_condition/>
<target_range>E35 Title</target_range>
<target_range_condition/>
</range_map>
<path_map>
<src_path_condition/>
```

```
<src_path>has</src_path>
<target_path_condition/>
<target_path>
<internal_link>P102 has title</internal_link>
</target_path>
</path_map>
</link_map>
<link_map>
<range_map>
<src_range>DC.TitleAlternative</src_range>
<src_range_condition/>
<target_range>E41 Appelation</target_range>
<target_range_condition/>
</range_map>
<path_map>
<src_path_condition/>
<src_path>has</src_path>
<target_path_condition/>
<target_path>
<internal_link>P132 has alternative form</internal_link>
</target_path>
</path_map>
</link_map>
</combined_links>
</map>
</mapping>
```

Figure2. A snapshot of the XML syntax of the crosswalk

Figure 3 describes the mapping of the DCCAP path DCCAP->marcel:OWN. This path corresponds to two equivalent CIDOC/CRM paths interlinked with an IsA relation: (a) E78(Collection)-*P24(transferred title of/changed ownership through)*-E8(Acquisition Event)-P14(carried out by/performed)- E39(Actor) and (b) the abbreviated form E78(Collection)-*P52(has current owner/is current owner)*-E39(Actor). In the case of the existence of more details about the collection owner, the CIDOC/CRM path: E78(Collection)-*P51 has former or current owner (is former or current owner of)*-E39(Actor)-*P76 has contact point (provides access to)*-E51(Contact Point) corresponds to the path DCCAP->cld:isAccessedVia, since the owner is considered to be the person or institution that holds the collection and allows the access to it for the audience.

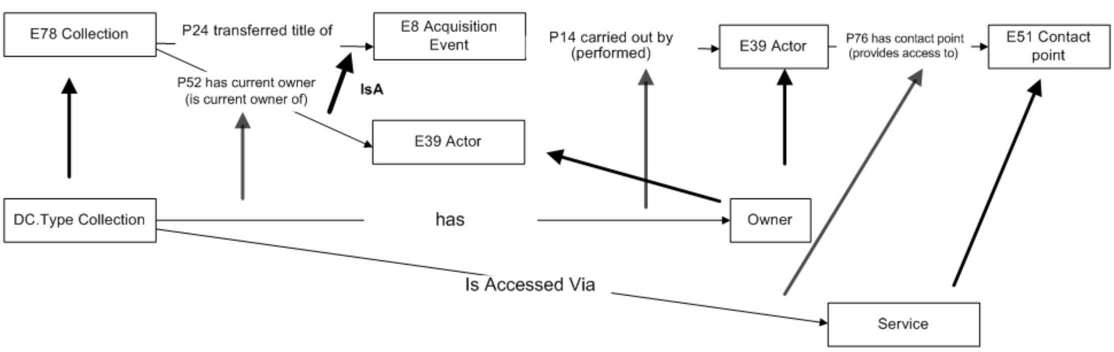

Figure 3. Mapping the collection owner to CIDOC/CRM

CONCLUSIONS

Metadata interoperability in cultural heritage domain is one of the main issues in the digital environment, due to the existing plethora of digital collections. In this paper we proposed a crosswalk between CIDOC/CRM and DCCAP (and vice versa) as a part of an integration mechanism that provides unified access to collection-level information. The most important prerequisite in this effort was to model cultural heritage collections as CIDOC/CRM paths. The main finding of the crosswalk development process is that many CIDOC/CRM paths that describe events taken place in the lifecycle of a collection, could not be mapped as DCCAP paths, since DCCAP does not provide relevant paths for them. For instance the CIDOC path: E78(Collection)-*P25 moved (moved by)*-E9(Move), that denotes the transfer of a physical object like the collection to another place, does not correspond to any path of the application profile. This phenomenon is quite reasonable since CIDOC/CRM is a rich ontology with general semantics related with cultural objects and covers all the facets of the life of a physical object. A significant conclusion is that core ontologies provide rich semantics and therefore should be preferred as mediating schemas.

The crosswalk between the two schemas adopts a path – oriented approach. In general, the followed methodology reveals explicitly all the concepts and activities related to cultural heritage collections lifecycle and therefore it is able to integrate a variety of cultural heritage application profiles.

REFERENCES

Gill T., Miller P. (2002), Re-inventing the Wheel? Standards, Interoperability and Digital Cultural Content. *D-Lib Magazine*, 8/1.

Dempsey L. (2000), Scientific, Industrial, and Cultural Heritage: a shared approach. A research framework for digital libraries, museums and archives. *Ariadne*, 22. Available from http://www.ariadne.ac.uk/issue22/dempsey/; accessed 18 May 2008.

Dublin Core Collection Description Task Group. (2006), *Dublin Core Collections Application Profile (DCCAP)*, Available from http://www.dublincore.org/groups/collections/collection-application-profile/; accessed 18 May 2008.

ICOM/CIDOC Documentation Standards Group AND CIDOC CRM Special Interest Group. (2006), *Definition of the CIDOC Conceptual Reference Model,* Available from http://cidoc.ics.forth.gr/; accessed 18 May 2008.

Gill, T. (2004), Building semantic bridges between museums, libraries and archives: the CIDOC Conceptual Reference Model. *First Monday*, 9/5. Available from http://www.firstmonday.org/Issues/issue9_5/gill/; accessed 18 May 2008.

Powell, A., *RSLP Collection Description Schema: Collection Description Schema*, Available from http://www.ukoln.ac.uk/metadata/rslp/schema/; accessed 18 May 2008.

ANSI/NISO. *Z39.91 - 200x Collection Description Specification – DSFTU*, Available from http://www.niso.org/topics/ccm/ccmstandards/; accessed 18 May 2008.

LC. (2002), *Encoded Archival Description, Version 2002*, Library of Congress. Available from http://www.loc.gov/ead/; accessed 18 May 2008.

Chapman A. (2004), Collection-level Description: joining up the domains. *Journal of the Society of Archivists*, 25/ 2, pp. 149-155.

Hakala, J. (2005). Collection mapping: models and standards for international cooperation. *In Proceedings of the International Seminar on Collection Mapping*, Helsinki, Finland.

Hunter J. (2001), MetaNet — A Metadata Term Thesaurus to Enable Semantic Interoperability Between Metadata Domains. *Journal of Digital Information*, 1/8. Available from http://jodi.tamu.edu/Articles/v01/i08/Hunter/; accessed 18 May 2008.

Pierre ST. M. and LaPlant, W. P. Jr. (1998), Issues in Crosswalking Content Metadata Standards, NISO White Papers, Available from http://www.niso.org/publications/white_papers/crosswalk/; accessed 18 May 2008.

Woodley M. (2000), Crosswalks: the path to universal access. In: Introduction to metadata: pathways to digital information. Online edition version 2.1. Available from http://www.getty.edu/research/conducting_research/standards/intrometadata/path.html ; accessed 18 May 2008.

MIT LIBRARIES. (2007), Metadata Reference Guide, http://libraries.mit.edu/guides/subjects/metadata/mappings.html

Day M. (1996), Metadata: Mapping between metadata formats, UKOLN: The UK Office for Library and Information Networking, Available from http://www.ukoln.ac.uk/metadata/interoperability/; accessed 18 May 2008.

Cruz I.F. and Xiao H. (2005), The Role of Ontologies in Data Integration. *Journal of Engineering Intelligent Systems*, 13/4.

DARPA. (1990), *Knowledge Sharing Effort (KSE)*, Defense Advanced Research Projects Agency. Available from http://www-ksl.stanford.edu/knowledge-sharing/; accessed 18 May 2008.

NamY-K., Coguen J., Wang G. (2002). A metadata integration assistant generator for heterogeneous distributed databases. In Robert Meersman and Zahir Tari, (eds.), *Papers from the International Conference on Ontologies, DataBases, and Applications of Semantics for Large Scale Information Systems*, vol. 2519, pp 1332-1344.

Tous R. Delgado J. (2006), Contorsion: A Semantic XPath Processor. In Elsevier (eds.), *Electronic Notes in theoretical Computer Science, Papers from the International Workshop on Database Interoperability (InterDB 2005)*, Namur, Belgium, pp. 87-102.

Amann B., Fundulaki I., Scholl M., Beeri C. Vercoustre A. M. (2001), Mapping XML Fragments to Community Web Ontologies. In Giansalvatore Mecca, Jérôme Siméon (eds.), *Proceedings of the Fourth International Workshop on the Web and Databases*, Santa Barbara, California, USA, pp. 97-102.

Hong B., Hongzhe L., Jiehua Y., Hongwei X. (2005), An Ontology-Based Semantic Integration for Digital Museums. In Springer (eds.), *Lecture Notes in Computer Science, Papers from the 6th International Conference on Web-Age Information Management*, Hangzhou, China, pp. 626-631.

Crofts N., Doerr M., Gill T. (2003), The CIDOC Conceptual Reference Model: a Standard for Communicating Cultural Contents. *Cultivate Interactive,* 9. Available from http://www.cultivate-int.org/issue9/chios/; accessed 18 May 2008

Hunter J. (2002), Combining the CIDOC CRM and MPEG-7 to Describe Multimedia in Museums. In D. Bearman and J. Trant (eds.), *Proceedings of the 6th International Conference Museums and the Web*, Boston, USA.

Doerr M. (2000), Mapping of the Dublin Core Metadata Element Set to the CIDOC CRM. Technical Report 274, ICS-FORTH, Greece. Available from http://cidoc.ics.forth.gr/crm_mappings.html; accessed 18 May 2008.

Kakali C., Lourdi I., Stasinopoulou T., Bountouri L., Papatheodorou C., Doerr M., Gergatsoulis M. (2007), Integrating Dublin Core metadata for cultural heritage collections using ontologies. ACM (eds.), *Proceedings of International Conference on Dublin Core and Metadata Applications (DC 2007)*, Singapore, pp. 128-139.

Doerr M., LeBoeuf P. (2007), Modelling Intellectual Processes: The FRBR - CRM Harmonization. *Digital Libraries: Research and Development*, Springer Berlin/ Heidelberg, pp.114-123.

Doerr M. (2006), Increasing the Power of Semantic Interoperability for the European Library, *ERCIM News*, 66, pp. 26-27.

Kondylakis H., Doerr M., Plexousakis D. (2006), Mapping Language for Information Integration. Technical Report 385, ICS-FORTH, Greece. Available from http://cidoc.ics.forth.gr/docs/Mapping_TR385_December06.pdf; accessed 18 May 2008.

DCMI. (2008), Dublin Core Metadata Element Set, Version 1.1, Dublin Core Metadata Initiative. Available from http://purl.org/dc/documents/rec-dces-19990702.htm; accessed 18 May 2008.

DCMI. (2001), Dublin Core Metadata Glossary, Dublin Core Metadata Initiative. Available from http://library.csun.edu/mwoodley/dublincoreglossary.html; accessed 18 May 2008.