European standards for the documentation of historic buildings and their relationship with CIDOC-CRM

Paola Ronzino[1], Nicola Amico[1], Achille Felicetti[1], Franco Niccolucci[2]

[1]PIN, VAST-LAB, Prato, Italy
{paola.ronzino, nicola.amico, achille.felicetti}@pin.unifi.it
[2]PIN, VAST-LAB, Prato, Italy
franco.niccolucci@unifi.it

Abstract. Integration of architectural datasets concerning historic buildings depends on their interoperability, which has as first step a mapping to a common schema. The paper investigates current approaches and proposes mapping to a CIDOC-CRM extension as the common glue to overcome the fragmentation of datasets provided by large national institutions such as MIBAC in Italy, EH in the UK, and so on, and by EU projects, each one structured according to a different metadata schema. The paper describes the mapping of the MA-CA MIBAC-ICCD schemas, probably the most comprehensive, to CRM.

Keywords. CIDOC-CRM, historic buildings

1 Introduction

There is a clear need in Europe of harmonizing actions on built heritage to face the challenges posed by environmental hazards and societal changes. The most comprehensive initiative on this regard is the EU Joint Programming Initiative on Cultural Heritage and Global Change [1], a framework within which EU Member States jointly address areas where public research programmes can respond to major societal challenges concerning heritage and its preservation. The theme has been addressed also by the EU project EU-CHIC (Cultural Heritage Identity Card) [2], which defined the concept of the CHICEBERG Protocol for the integrated documentation of built heritage, based on a taxonomy of historic buildings developed by the EU project Perpetuate [3]. EU-CHIC mainly concerns the conservation and documentation of environmental changes affecting built heritage assets, such as historic buildings and monuments. Most countries in Europe have developed their own systems for storing information concerning built heritage: among others, the Italian Ministry of Culture MIBAC that adopts forms prepared by a specialized institute (ICCD, Central Institute for Cataloguing and Documentation [4]); English Heritage, using the MIDAS scheme [5]; the French Ministère de la Culture, using the Schéma Documentaire Appliqué au Patrimoine et à l'Ar-

chitecture (SDAPA) [6]. Moreover, European projects contributing to Europeana, the European digital Library, have developed their own schemas and mapped them to EDM, the Europeana Data Model. Such projects include CARARE [7] and 3D ICONS [8]. In conclusion, there is a number of different metadata schemas organizing large datasets but preventing any effort for dataset integration, which is an absolute need to develop European policies for research, conservation, restoration and dissemination. Such datasets intersect those considered by ARIADNE [9], the European Research Infrastructure for archaeological datasets, as far as built heritage includes archaeological remains. ARIADNE aims at providing an integrated access to archaeological datasets throughout Europe, and is developing an extension of CIDOC-CRM to guarantee their interoperability [10]. It seems therefore that CIDOC-CRM, or if necessary an extension of it, is the key to overcome the fragmentation of architectural datasets, and this is the way we propose to follow. We are currently building a mapping from each of the metadata schemas used in the most important European repositories, such as those mentioned above, i.e. the ICCD schemas, MIDAS, CHICEBERG and the CARARE/3D ICONS schemas, to the CIDOC CRM. It is a complicated work, because it involves more than 700 fields, some identical in meaning, some just similar but with a different nuance, and other very different. A preliminary version of the mapping is ready and will be published on VAST-LAB's web site [11]. The mapping of the CARARE schema to CIDOC CRM has been discussed in [12].
In our experience, the most comprehensive is the ICCD one, and we are working closely with the Institute to develop the mapping of the many forms it uses. A full description of the forms may be found on the ICCD site [4].In this paper we will present a draft mapping of the ICCD Monument form to CIDOC CRM; or, better, an outline of it, for space reasons. The full version is going to be available on the above-mentioned VAST-LAB's web site as well.

2 The ICCD MA/CA form

The MA/CA form is used for archaeological monuments and complexes [13]. As regards architecture, there is a similar form called form A [14], used for historic buildings, which has only slight differences from MA/CA. We have mapped both, but for the sake of brevity we will present here only the MA/CA to CRM mapping. The MA/CA form includes more than 300 fields, each identified by a unique letter code and a name. We will use only the code and give an informal English translation of the name. Metadata are grouped in the following 'wrappers': CD-AC – Codes; RV – Relationships; OG – Object; LC – Current Location; CS – Cadaster; LS – Historic Location; GP-GL-GA – Georeferencing; RE – Way of discovery; DT – Chronology; AU – Cultural definition; RO – Reuse; MT – Technical data; CO – Conservation; RS – Res-

toration; DA – Analytical Data; MC – Samples and analyses; TU – Legal status; DO – Sources; AD – Data access; CM - Compiler; AN – Notes.
Fields (and wrappers) of little interest for integration will not be considered.

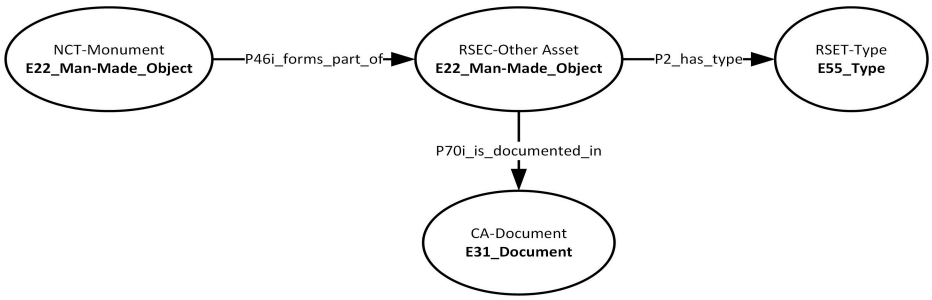
3 The mapping

3.1 RV – Relationships

This set of fields is used to document the relationship of the monument, identified with its unique code NCT, with other assets of different kind. In the relationships below, the domain is the monument and the range is the other asset, which can belong to the same category or can be different. Entities corresponding to MA/CA fields are identified with the MA/CA letter code.

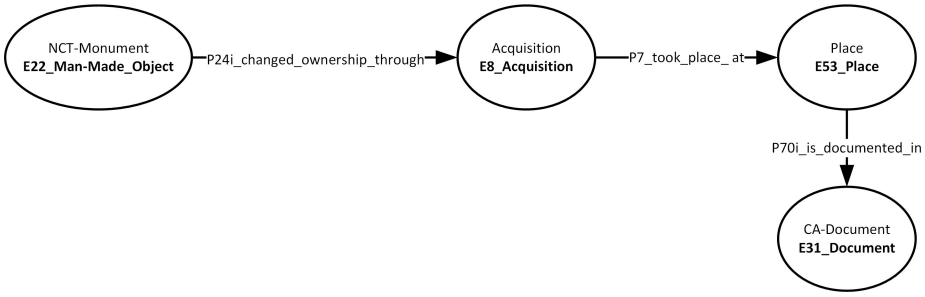
- Is contained in:

The monument relates to another monument (MA) or archeological complex (CA), which represents the monument location at the time of cataloguing.



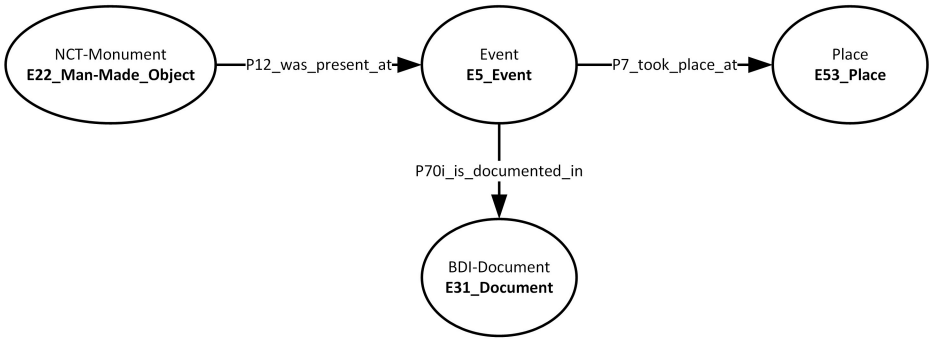
- Was found in:

This relation links the monument (MA) or archaeological complex (CA) to the site (SI form) or Stratigraphic Essay (SAS form) where it was found.



This path is not completely convincing and perhaps a better way of documenting archaeological discovery could be considered in a future extension of CIDOC CRM.

- Is involved in:

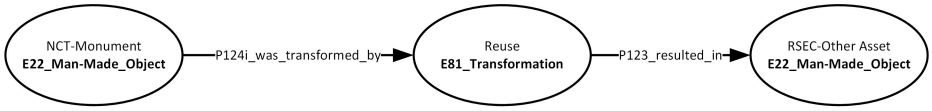


This documents the connection between the monument, and an event (such as a festivity, celebration, rite, etc.), documented in a form pertaining to intangible heritage.

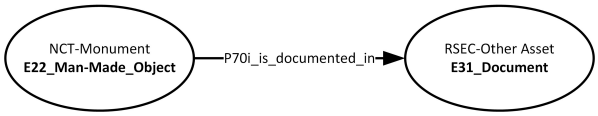
- Has environmental/spatial relationships with:

- Was made in:

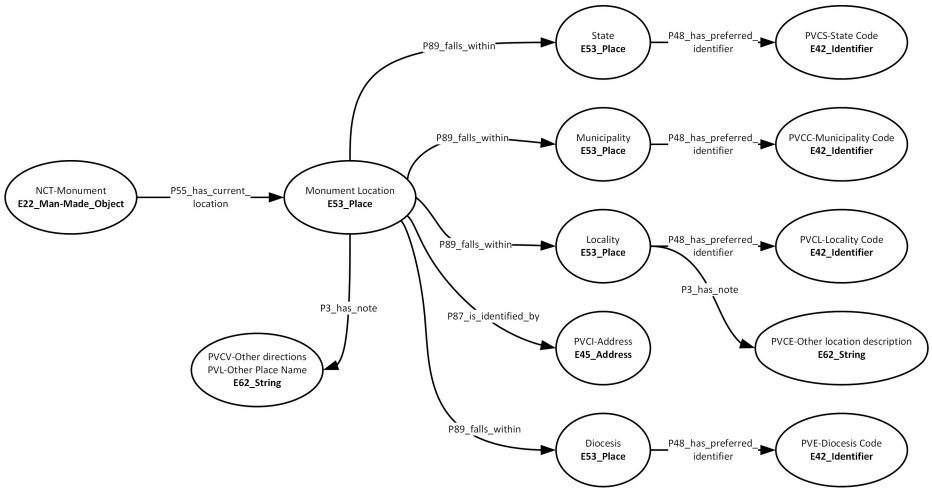
- Is reused by:



- Is documented in:



3.2 LC – Current Location

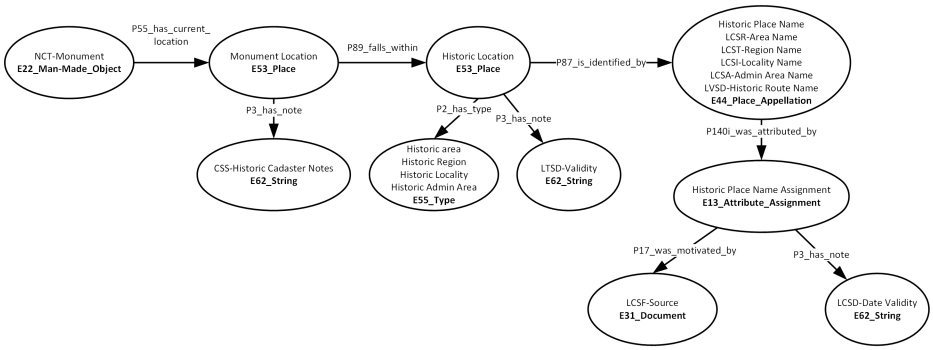


As shown by the diagram above, metadata about location are modeled via the monument location (E53) that falls within (P89) various other places useful to define the location.

3.3 LS – Historic Location

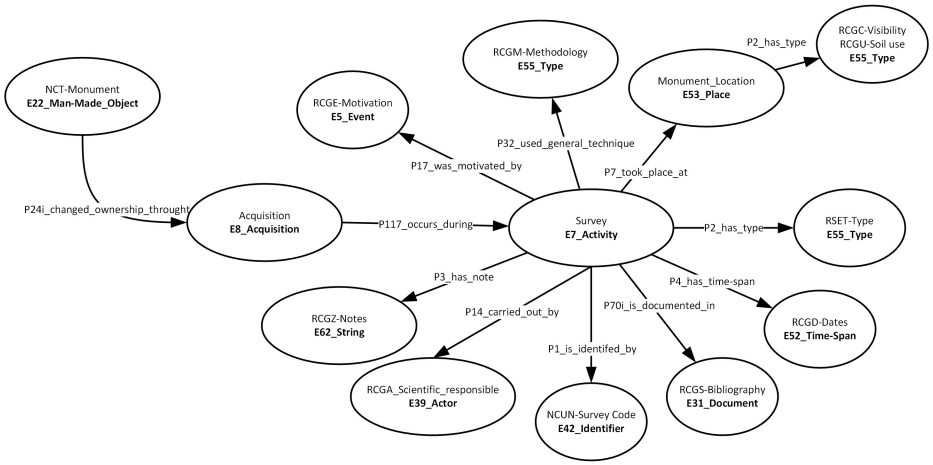
Historic Location relates the monument to various historic places, such as areas, roads and places with their place names. This correspondence is modeled via the Monument Location, as before, which receives (P140i) by an Attribute Assignment (E13) the assignment of various historic locations (E53) with their place names or other specification (E44 Place Appellation).

We used an E62 String to express the time validity of the historic reference as a note to the Historic Place Name assignment, since CIDOC-CRM does not seem to have a simple way of expressing the time validity of a historic localization.

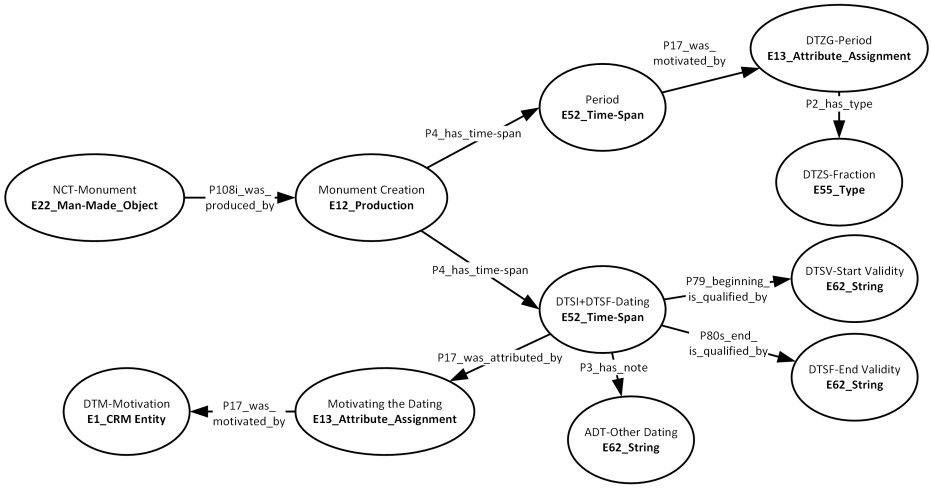


3.4 RE – Way of Discovery

This wrapper collects information about the way the monument was discovered, distinguishing among survey, excavation and other investigations. The diagram below concerns the survey, while the excavation one is very similar. The modeling starts with an 'Archaeological Discovery', on which the same comments as above can be made. In this case it occurred during a Survey (E7) Activity, identified by its code NCUN for which – as for any field whose code begins with N – there is an authority file. The Survey took place (P7) at the Monument Location (E53) about which RGCU Soil Use and RCGC Visibility of the terrain are recorded as types (E55). Information about the survey concerns among others its RCGD Date (E52 Time Span), RCGA who did it (E39 Actor), and the Methodology type (E55) used. The reason RCGE for carrying out the survey is modeled as an E5 Event.



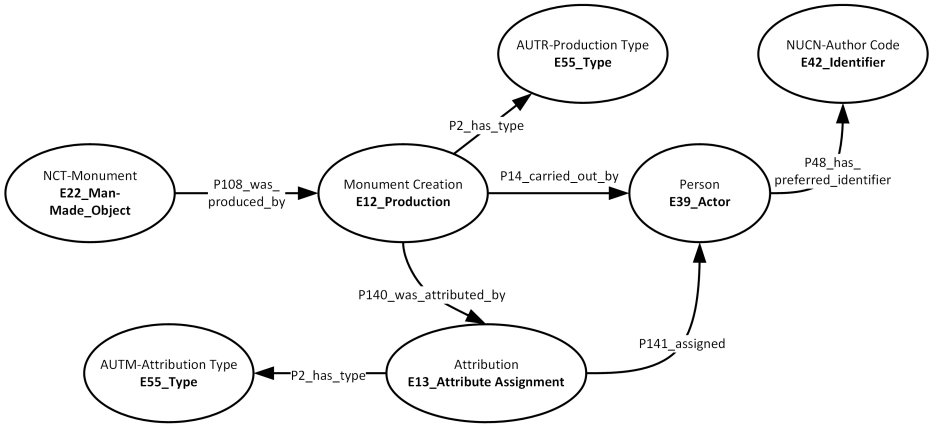
3.5 DT – Chronology



The chronology section is based on an E12 Production event. Chronology may be approximate, falling within the DTZG Period (E52 Time-Span), affected by a qualifier DTZS Fraction, modeled as E55 Type, e.g. 'end of', 'early', and so on; or more precise, but possibly still approximate such as "ante 1410 AD", "approx. 600 BC" etc., with a start and an end date incorporated in DTSI+DTSF Dating, an E52 Time-Span, start qualified (P79) and end qualified (P80) by validity, respectively DTSV and DTSL, as 'ante', 'approx.' etc.

3.6 AU – Cultural Definition

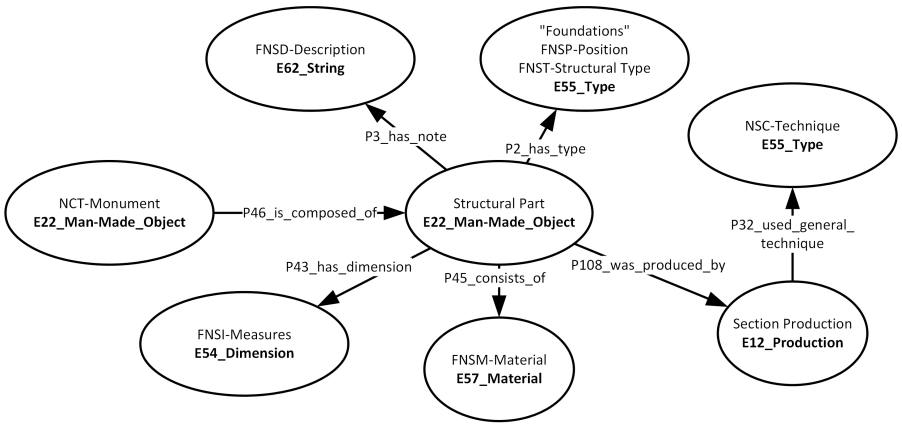
This section concerns authorship, and is centered on the Monument Creation, an E12 Production. The Author is an E39 Actor. It could be an E21 Person identified (P48) by the NUCN Author Code that refers to the AUT authority file, which includes all the information concerning the author. If the identification is imprecise, reference to "school of", "workshop of", or "group of" is included in a special field called AUTS. These special cases lead to slightly different modeling (not presented here for space reasons), where the Author is an E74 Group and the participation of a person in this is modeled with P15 was influenced by, for "school of"; P107i is current or former member of, for "group of"; and so on. AUTM, the Motivation of the attribution, is modeled via an E13 Attribute Assignment, which assigns the Author to the Production. The mapping of additional information, sometimes present, concerning the cultural ambit, i.e. generic cultural references to a cultural context, and the commission of the monument is not detailed here for the sake of space.



3.7 DA – Analytical Data

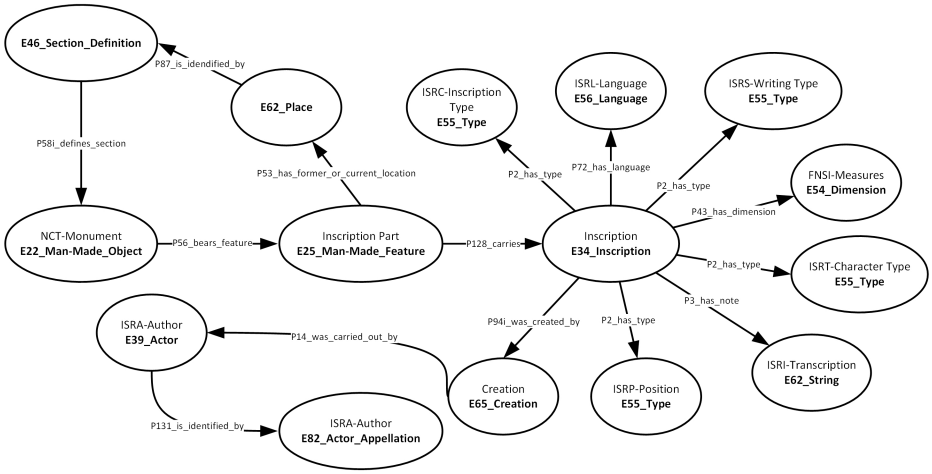
This section describes the structural parts of the monument: foundations, vertical and horizontal structures, stairs, the roof, open spaces, and includes marks, inscriptions and emblems. Each one of these has a separate subsection.

The diagram below concerns foundations. They are modeled as a part of the monument, defined as another E22 Man-Made Object of type “Foundations”. Besides the FNSD Description, modeled as an E62 String, and several types assigned to the part, the information recorded includes FNSM Material, modeled as E57 Material, the material used for the foundations such as bricks, stones, unknown etc.; and the construction technique, modeled via an E12 Production event relating to the part, which used as general technique (P32) the FNSC Technique, an E55 Type. Information concerning horizontal and vertical structures, the stairs, the roof and open spaces is very similar and is modeled in the same way, with more types characterizing the different parts.



3.8 ISR – Inscriptions

The following diagram describes the model for the inscriptions.



The interpretation of the modeling is straightforward. A difficulty here concerns the text and author of the text of the inscription. In some cases the original ISRA Author field contains mixed information, such as the author and the work from which the inscription text is taken, so modeling it as an E62 String is somehow compulsory, as a consequence of overloading the field with too much information in the source data model. But in other cases, if for example only the text author is documented and further elaborated with information on the person, modeling it as a String leads to a cul-de-sac. To provide a more structured and detailed information, whenever possible both author identification and attribution must be described. To identify the author, a path such as E34 Inscription – P94i was created by – E65 Creation – P14 carried out by – E39 Actor – Actor P131 is identified by – E82 Actor Appellation, may be used. If comments on the attribution must be made, e.g. to qualify its reliability or source, this path may be substituted with E34 Inscription – P140 was attributed by – E13 Attribute Assignment (of authorship) – P140 assigned – E39 Actor, and then further qualifying the authorship attribution E13.

4 Conclusions and Further Work

For space reasons, it was impossible to present here a complete description of the mapping, but we hope that the section dealt with gave the flavor of the work. In conclusion, the mapping is feasible and perhaps improves the original documentation scheme without loosing its richness of details. The ongoing mappings of other national repositories of monument documentation, and the

creation of multilingual thesauri that are also in progress (see [15] for further details) show that the interoperability of monument datasets is feasible, if not easy, and that the integration of these repositories at European level would create an infrastructure as useful as the forthcoming archaeological one. Further work will concern completing the mapping of other ICCD schemas relating to architecture and addressing conservation and restoration, which are present in these forms in a very succinct way.

5 Acknowledgements

The present work has been partially supported by the European Union through the ARIADNE project.

References

1. JPI: http://www.jpi-culturalheritage.eu
2. CHICEBERG http://eu-chic.eu/index.php/news/entry/chiceberg/
3. Perpetuate: http://www.perpetuate.eu
4. ICCD: http://www.iccd.beniculturali.it
5. MIDAS: http://www.english-heritage.org.uk/publications/midas-heritage/
6. SDAPA: http://www.culture.gouv.fr/culture/dp/schemaDAPA/index.html
7. CARARE metadata schema 2.0, http://www.carare.eu/eng/Resources/CARARE-Documentation/CARARE-2.0-schema
8. 3D ICONS: http://www.3dicons-project.eu
9. ARIADNE: http://www.ariadne-infrastructure.eu
10. Ronzino, P., Niccolucci, F., D'Andrea, A.: Built Heritage metadata schemas and the integration of architectural datasets using CIDOC-CRM. Paper accepted at "Built Heritage" Conference, Milan November 2013
11. Ronzino, P., Amico, N., Felicetti, A., Niccolucci, F.: Built heritage documentation in CIDOC-CRM. PIN technical report http://www.vast-lab.org/documents (2013)
12. Ronzino, P., Amico, N., Niccolucci, F.: Assessment and comparison of metadata schemas for architectural heritage. In Proc. of the XXIII International CIPA Symposium (2011)
13. ICCD Scheda MA-CA Monumento Archeologico Complesso Archeologico v.3.00 (2013) http://www.iccd.beniculturali.it/index.php?it/251/beni-archeologici
14. ICCD Scheda A Beni architettonici e ambientali Edifici e manufatti v.3.00 (2013) http://www.iccd.beniculturali.it/index.php?it/252/beni-architettonici-e-paesaggistici
15. Ronzino P., Amico N., Niccolucci F. Federating Specialized Digital Libraries. In: Niccolucci F., Dellepiane M., Pena Serna S., Rushmeier H., Van Gool L. (eds) Proceedings VAST2011, Eurographics, pp. 97-103