

Zeri e LOD.

Extracting the Zeri photo archive to Linked Open Data: formalizing the conceptual model

Ciro Mattia Gonano
Dept. of Computer Science
and Engineering
University of Bologna
Mura Anteo Zamboni 7
40126 – Bologna (Italy)
ciromattia@gmail.com

Francesca Tomasi
Dept. of Classical Philology
and Italian Studies
University of Bologna
via Zamboni 32
40126 Bologna (Italy)
francesca.tomasi@unibo.it

Francesca Mambelli
Fondazione Zeri
University of Bologna
Convento di Santa Cristina
Piazzetta Giorgio Morandi 2
40125 Bologna (Italy)
francesca.mambelli6@unibo.it

Fabio Vitali
Dept. of Computer Science
and Engineering
University of Bologna
Mura Anteo Zamboni 7
40126 – Bologna (Italy)
fabio@cs.unibo.it

Silvio Peroni
Dept. of Computer Science
and Engineering
University of Bologna
Mura Anteo Zamboni 7
40126 – Bologna (Italy)
silvio.peroni@unibo.it

ABSTRACT

This paper presents the first steps of a project to convert the notable Italian “Zeri photo archive” to a linked and open dataset. The full project entails the analysis of the records’ description model (*Scheda F*) in order to define a suitable ontology by exploring existing data models, the creation of the RDF triple store, the creation of links to the cloud, and the definition of the user interface for browsing the linked open dataset. This paper presents and discusses the conceptual modeling of the data stored in the Zeri archival database.

Categories and Subject Descriptors

H.3 [INFORMATION STORAGE AND RETRIEVAL]: H.3.3 Information Search and Retrieval, H.3.5 Online Information Services, H.3.7 Digital Libraries; [DOCUMENT AND TEXT PROCESSING]: 1.7.2 Document Preparation; J.5 [ARTS AND HUMANITIES].

Keywords

FRBR, FEO, Scheda F, CIDOC-CRM, PROV-O, FABIO, RDF, OWL 2 DL.

1. INTRODUCTION

Employing Linked Open Data (LOD) concepts in the cultural heritage domain (CH) is a challenging field of study. We are not the first to notice that “libraries, museums, and archives [LAM]

are figuring out new ways to export their data in triples, integrate external linked datasets into their collections, and develop new interfaces for users to experience cultural heritage” [5].

Even if plenty of activities in the production, integration and dissemination of LOD can already be noticed¹, much still has to be done. First of all, plenty of CH data is still silo-ed in traditional ways and has yet to be transformed in open datasets. Secondly, the conversion procedure is not culturally neutral, because it involves choices on both the metadata vocabulary describing the collection (the *description model*) and the ontology used for the conversion process (the *conceptual model*).

The heterogeneity of the standards used by the institutions devoted to the preservation of CH data is a delicate question [20], and the solutions adopted, and the different ontologies used to expose these data is a preeminent topic as well (e.g. EDM [4], [6]). Mapping ontologies to the LOD domain (see. e.g. [2]) is far from a trivial task, not only from a technical point of view but also and predominantly from a conceptual point of view.

Populating the LOD cloud requires mechanisms, that automatize this operation for the huge data sets we own, but also and most importantly a clear understanding on what the Web already provides. The growth of the LOD cloud in these last years (see the diagram in [12]) requires one to deeply understand the possibilities to use fruitfully in the CH domain the information in the Web of data already existing (a survey can be found in [5]). Catalogues and inventories of LAM are gold mines but their thorough and fruitful exploitation is only possible by the appropriate use of sophisticated Semantic Web methodologies and tools. For instance, conversion projects could choose to proceed in two directions: either to produce a direct, flat representation of the existing descriptive data of their treasures (in which every item of the collection is seen as one ontological individual, collapsing all properties onto itself) or to develop a subtler, more nuanced approach based on a layered separation of concerns, in which

¹ E.g. LOD LAM activities, *Linked Open Data in Libraries, Archives, and Museums*: <http://lodlam.net>

every item is represented as a set of layered concepts with clearly distinguished concerns and characterizations. In [18] we already discussed the representational chaos that derives from collapsing information layers, in that case information about the *item* being described (i.e., a b/w photography of the '60) with information about the *subject* of the item (i.e., a 18th century lithography of the Coliseum in Rome).

In this paper we present *Zeri e LODE*, a contribution towards a properly layered characterization of cultural items, where such representational chaos is absent or strongly reduced. Our starting point is a model case of this situation: The Zeri Photo Archive, an archival collection of photographs of paintings described in a traditional RDBMS that we want to convert to LOD by adopting a layered conceptualization (namely, CIDOC-CRM) as both a descriptive and a conceptual model.

We first proceeded to reengineer the Entity/Relationship model provided by the database tables, which structures data according to the *Scheda F* (Italian for *F entry*, a description standard issued by the Central Institute for Cataloguing and Documentation [ICCD] for the cataloguing of photographic materials² – where *F* stands for “Fotografia”, photography in Italian), into OWL, so as to obtain a first version of an ontology, that we call *FEO*, *F Entry Ontology*.

This is the first phase of a complete reconversion project, that will see the transformation of the data currently stored in the database into RDF statements compliant to the OWL ontology we are developing, and the use of automatic and semi-automatic tools to generate links to existing datasets. The ontology itself is being iteratively enhanced following the modifications of the specifications of the ICCD *Scheda F* and of CIDOC-CRM, making sure that the whole conceptual organization and entity naming of the existing model are affected as little as possible.

FRBR, in particular in its second release commonly known as FRBRoo³, is a characterization of intellectual productions such as texts, arts and photographs, and it is an important part of the CIDOC-CRM conceptual model. Our mapping process adopts it as a model to follow for both the implementation of the ontology, based on common model (mostly FABIO and Prov-O) and the mapping between the ICCD *Scheda F* and the CIDOC-CRM.

The paper is organized as following: in section 2 we describe the database; in section 3 we give all the appropriate bibliographic context; in section 4 we present the project; in section 5 we focus on the first steps of the project; in section 6 we present the future work involved in the project.

Although in this paper we present the whole project, we will particularly emphasize the steps regarding the modeling of the *Scheda F*, the creation of the ontology creation and the mapping between the *Scheda F* and the CIDOC-CRM ontology.

2. THE ZERI PHOTO ARCHIVE

The photographic library bequeathed by Federico Zeri (1921-1998) to the University of Bologna is a collection of 290,000 photographs of artworks and monuments, organized, expanded and improved by the scholar throughout his entire career.

Similarly to other famous international art historians and *connoisseurs*, like Bernard Berenson and Roberto Longhi, Federico Zeri considered the photo archive a fundamental tool for studying as well as for philological analysis of the artworks. He worked in direct contact with the photographs which he constantly moved around his desk, combining details or subjects and

recreating sequences. This method relied on a sensitive eye, trained to notice colour nuances apparent on b/w prints and, paradoxically, blurred and distorted on colour prints.

The uniqueness of this collection lies in the rarity of the materials and the documentation of a wide range of artworks. Many ancient photos (approx. 40% of the records dates from before the end of the 1930s), often documenting artworks in different moments of their history, different states of preservation or before restorations, can only be found in Zeri's Library. The photographs come from national photo archives (Fototeca Nazionale, Soprintendenze), Italian and foreign museums, private archives and collections, auction houses, restorers. Some were purchased by the scholar or donated to him. A complex cross-reference system, in the form of annotations on the back of the photos, provides information about provenance, attributions, restoration and bibliography of the artworks and connects the photographic library to the art library (46,000 books) and the auction catalogue collection (37,000 catalogues, the most important in Italy). Moreover, the personal arrangement of the photo archive reflects Zeri's concept of art history and transmits his attributions.

Since its constitution in 2002, Federico Zeri Foundation's first task was to ensure that the Photographic Library was made accessible to scholars and students by cataloguing and digitizing the original photographs and setting up a freely available online database.

The cataloguing of the phototypes was preceded by a phase of theoretical development. It was necessary to adopt an entry that was compatible with national catalographic standards issued by the Italian Cultural Heritage Ministry, but also structured to respect the specificity and organisation of the Zeri Library. Given the main subject and destination of the archive, it was compulsory to take into consideration also future users. The aim was to create a computerised catalogue of photographs (the asset physically possessed) for consultation mainly by art historians, whose approach would, in the majority of cases, be based on details of the works recorded.

The need to save and standardise all the levels of information on the back of the photographs, combined with that of conveying the hierarchical structure of the Zeri Library to those consulting it, required an expansion of ministerial entries.

A record has been therefore created comprising two primary catalogue units, one describing the photograph and the other the describing the work of art represented in it. These two entries and their connections give rise to a symmetrical database that allows two-sided research by users and transversal navigation.

The data model adopted for photograph description is the *Scheda F* of the Central Institute for Cataloging and Documentation (ICCD). Of this standard, composed in its extended version by more than 300 fields, the Zeri Foundation uses only a subset of 113 fields organized in blocks, called *paragraphs*, mimicking the organization of the data on paper and referring mainly to the inventory level: photographer, chronology of the positive, subject, title, location and classification, technical data, state of preservation, copyright, codes.

In addition, other paragraphs are specified where to record information referring to the catalogue-level. The “Place and date of the shot” paragraph allows art historians to date a certain state of preservation documented by the picture; the “Previous locations” paragraph can lead to the virtual reconstruction of ancient photo archives dismembered by Zeri in his collection; the “Editing and publishing” paragraph refers to other authorships involved in the production or re-use of the phototypes (commissioners, editors, printers).

² <http://www.iccd.beniculturali.it/getFile.php?id=1185>

³ http://www.cidoc-crm.org/frbr_inro.html

The *Scheda OA* (*Artwork entry* – Opera d'Arte in Italian – derived from the ministerial model from which it draws inspiration) refers to the semantic content of the photography. It is composed of 79 fields that offer a detailed description with related search prompts such as type, subject, attributions, chronology, current and previous locations, material and technique, reference bibliography, exhibitions and attached documentation. This information is deduced primarily from Zeri's annotations on the back of the photos, as well as from the documentation tied to it and from volumes available in the library.

Each *Scheda OA* is linked to all *Scheda F* of its photographs. These are illustrated by digital images in a variety of formats (such as low and high resolutions, thumbnails and scanned backs of the photos).

Six Authority Files enrich the database providing special data-sets about photographers, archival description of the collection, artists, bibliography, auction catalogues, and associated documents. Some of these authorities (Associated Documents, Auction Catalogues) are accessible on line as autonomous databases.

The cataloguing, which until now has extended to 160,000 photographs covering the main sections of the archive, *Italian Paintings*, *Italian Sculptures* and *Still Life*, has led to the creation of a freely available online database accessible at <http://www.fondazionezeri.unibo.it>. The database has been consulted by more than 70,000 individual users in 2013 and has become an essential point of reference for the international community of art historians, so much so that it was described as "The first, and no doubt the most important database on art history" (*The Art Tribune*, 16/8/10). Such recognition has been achieved thanks not only to the quality and rarity of the images, or to the user-friendly nature of the archive, but also to the analytical nature and scientific reliability of the metadata of both the photographs and their iconographic contents.

The exportability of the data in xml format and mapping carried out on other data models has allowed the Zeri Foundation to contribute as content provider to portals and international projects, such as Europeana and Artstor⁴.

Recent developments in the project have led to extending the cataloguing to other types of documentary assets and to a dialogue between different cataloguing systems, all within one single context. The indexing now includes documents traced among the photographs in the archive and the Foundation's auction catalogues. Each item, described according to archival and library sciences rules, is imported into the platform and linked to the artwork entry it derives from. The artwork entry therefore serves as a core aggregator of information obtained from historic and artistic research sources collected and used by Federico Zeri: photographs, books and paper documents.

3. RELATED WORKS

There are many projects regarding different aspects of the transformation of CH collections into LOD, such as [3], [4], [6], [14], principally affecting the conversion of the database, the creation of the triple store, the link to the cloud, the browsing interface. Yet, the main focus of this paper is on the ontology and the main characteristics of the *Scheda F*, the CIDOC-CRM ontology, the role of FRBR.

The *Scheda F* standard was issued in 1999 after a long gestation. Its publication coincided with the belated recognition of photography as a cultural asset by the Italian legislation. Only

with the act "Testo Unico delle disposizioni legislative in materia di Beni Culturali e Ambientali" (DL n. 490/1999), photographs have been explicitly included in the list of objects worthy of national policies for the protection, conservation and exploitation. The photographs have been therefore promoted from the pure documentary function that was previously assigned to them.

Since then, the success of the *Scheda F* has been neither quick nor widespread. The capillarity and analyticity of the standard, especially in its extended version (catalogue-level) and the complexity of the filing rules that combine cataloguing science with principles from disciplines such as archival and library sciences, worked as a considerable obstacle to its success. Despite these problems, that resulted in a dilation of time and cost of compiling, in about 15 years of application in Italy the *Scheda F* showed its scientific validity. In 2004, the standard was submitted to a non-substantial revision (*Scheda F* version 3.00), while a trial is currently underway that will lead to a more significant change, offering the possibility to describe complex sets, in addition to individual items (*Scheda F* 4.00).

Among the databases available on the Web and using *Scheda F*, we can highlight the following: the Regional Information System for the Friuli-Venezia Giulia Cultural Heritage (SIRPAC)⁵; the Artistic Heritage portal of Regione Lombardia⁶; the Guarini information system of Regione Piemonte⁷; the Regional Cultural Heritage databases of Regione Veneto⁸; l'Album di Roma: 20th Century private photographs⁹. The *Scheda F* produced by the State and Regional Authorities will all be available for consultation in the SIGECweb system¹⁰ once it will be released to the public.

Many institutions, though, use simplified versions of the *Scheda F*, adopting its fundamental data set, but compiling a much lower number of fields compared to those meant for inventory-level (minimal absolute requirement).

The complexity of the *Scheda F* makes it difficult to export it as a model. Nevertheless, the principles on which the *Scheda F* is set seem to have anticipated a trend that is now spreading among many institutions owning collections of photographs of art/historical documentation: an increasing interest for the photographic object itself, promoted from a mere documentary function and from its total subsidiarity to the work of art depicted.

Many institutions whose standards (and therefore whose databases) have been so far totally biased towards the description of the iconographic content of the photographs, are changing their cataloguing rules and adding fields for the description of the photograph itself. This is the case, for instance, of the Berenson Library, an institution of the Harvard University¹¹, and of some German photo archives (Hertziana, Kunst Historisches Institute of Florence) that are part of the Bildarchiv Photo Marburg¹².

This renewed interest in the photographic object does not prevent the persistence, in some cases, of a certain ambiguity in the identification of the real object of cataloguing (as it happens, for instance, in Europeana [18]).

⁵ <http://www.sirpac-fvg.org/>

⁶ <http://www.lombardiabeniculturali.it/fotografie/>

⁷ <http://www.regione.piemonte.it/cultura/GPCWeb/zf/>

⁸ <http://catalogo.regione.veneto.it/beniculturali/>

⁹ <http://www.albumdiroma.it/sicap/opac.aspx?Web=ROMA>

¹⁰ <http://www.iccd.beniculturali.it/index.php?it/118/sistema-informativo-generale-del-catalogo-sigec>

¹¹ <http://itatti.harvard.edu/berenson-library/collections/fototeca-photograph-archive>

¹² <http://www.fotomarburg.de/>

⁴ <http://www.artstor.org>

The mapping method has been conducted on CIDOC-CRM [1]. Experiments in using CIDOC-CRM¹³ are interesting for our project. They represent useful cases to be analyzed and are important works in the field. CIDOC-CRM is the widest ontology used in the CH domain and represents a shared model that gives opportunity in semantic conceptualization of photographic materials in an FRBR perspective. Experiences in mapping (e.g. [4], [22]) have to be considered as well.

Some experiences in producing LOD in CH domain are an important reference: Library of Congress¹⁴, VIAF¹⁵, CulturalItalia¹⁶, Bibliothèque Nationale de France¹⁷, Deutsche National Bibliothek¹⁸, Archives HUB¹⁹.

Existing linked and open datasets give a collection of suitable data and a model. Also vocabularies such as the Art & Architecture Thesaurus by the Getty Research Institute²⁰ or Iconclass²¹ represent important related sources.

FRBR [9] is widely used in the CH domain but sometimes non in the most appropriate way. Even if it was born in the bibliographic domain, its conceptual model has been adopted in different fields as a conceptual framework.

The Zeri collection is already contained in Europeana and Europeana is moving towards a LOD creation. So we are surely in line with the guidelines from the data.europeana.eu project²². Although we do consider the EDM model [6] and the EDM Mapping Guidelines²³ we are defining our own model, giving a greater role to the FRBRoo part and considering other ontologies, such as PROV-O and FABIO, not yet known in the CH domain.

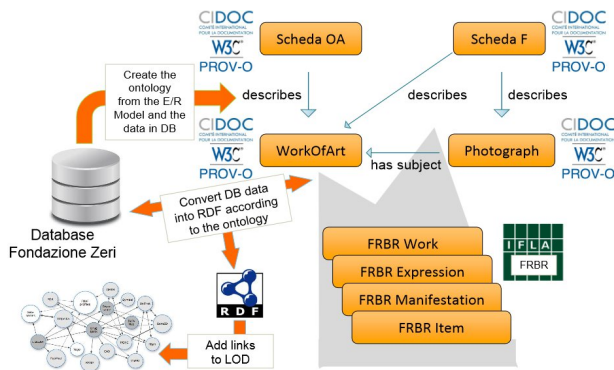


Figure 1. The steps of the project

4. THE ZERI E LOD PROJECT

The *Zeri e LOD* project consists in the steps in Figure 1:

1. the formalization of the *Scheda F* by mapping its schema onto CIDOC-CRM (section 4.1);
2. the creation of the triple store (section 4.2);
3. the adding of links to the LOD cloud (section 4.3);
4. the creation of the browsing interface (section 4.4).

4.1 Formalizing the *Scheda F*

In late January 2013, the Federico Zeri Foundation was invited together with the representatives of other thirteen important photo archives based in Europe and the United States, to a two-day colloquium in New York dedicated to the future perspectives for archives of art historical documentation. As a result of the meeting, the International digital photo archive consortium (IDPAC) was launched. One of the objectives of this group, which includes institutions with an aggregate 31 million photos of art, nearly all in the process of digitizing their archives, is to discuss the possibility, and verify the feasibility, of a common platform where to share all their digital resources and link all data for maximum accessibility, thereby creating a consolidated resource for art historical research.

This project, though is in its exploratory stages, identified CIDOC-CRM as the most suitable ontology to which to map data coming from multiple and dissimilar sources as it can represent most cultural heritage data standards.

In the conversion from the schema of the *Scheda F* to CIDOC-CRM we initially considered the specific flavor of the *Scheda F* used by the Zeri Foundation, with its subset of 113 fields (based on the experimental 1.04 version of the ICCD standard) and an handful of custom extensions to it.

The schema of the *Scheda F* (a sample in Figure 2), though impressively data-rich, has a flat structure for the documenting of the described object, and thus it lacks in flexibility on cross-linking related entities (for instance, many fields need controlled authority lists or dictionaries) and is rather error-prone for typos or misspelled words by human cataloguers.

Since exposing the data of the Zeri Photo Archive to the world is the main goal, we have chosen a standard such as CIDOC-CRM as the best way to access such data using a “common language”, so as to spread its availability to its maximum.

The actual mapping reflects a few trade-offs imposed by the very flexibility of CIDOC-CRM: its model is designed to cover many types of information models, from those used by museums to those used by libraries or by archiving organizations, so it lacks a few fields peculiar to photography, through which the *Scheda F*, being focused on them, deepens considerably its reach. Still, we tried to extend the CIDOC-CRM model as little as possible, so as to stick to the standard as far as we could (but see section 5.2 for a detailed description of this mapping).

¹³ http://www.cidoc-crm.org/uses_applications.html

¹⁴ Linked Data Service: <http://id.loc.gov>

¹⁵ <http://viaf.org/>

¹⁶ <http://dati.culturalitalia.it/>

¹⁷ <http://data.bnf.fr>

¹⁸ <http://www.dnb.de/EN/lds>

¹⁹ <http://datahub.io/dataset/archiveshub-linkeddata>

²⁰ <http://www.getty.edu/research/tools/vocabularies/aat/>

²¹ <http://www.iconclass.nl/home>

²² <http://pro.europeana.eu/linked-open-data>

²³ EDM documentation: <http://pro.europeana.eu/edm-documentation>



Figure 2. A portion of a record sample

4.2 Creating the RDF triple store

The triple store will be generated by converting the Zeri database into RDF according to the ontology described.

The idea is to work on the relational database in order to convert records into XML files (see e.g. [3]) and then to convert the assertion contained therein into RDF triples using XSLT.

A first step in this direction has been already made, as we already can access the database records in a reasonable and easy to work with XML format, as shown in Figure 3.

```
<PARAGRAFO etichetta="Autore">
  <AUFN etichetta="Autore della fotografia">Detroit Institute of Arts</AUFN>
  <AUIF etichetta="Indicazione del nome e dell'indirizzo">Photographic Dept. Detroit Institute of Arts</AUIF>
  <AUFM etichetta="Motivazione dell'attribuzione">Simbro</AUFM>
  <EOL etichetta=""/>
  <AUTN etichetta="Autore dell'opera fotografata">Giovanni del Biondo</AUTN>
</PARAGRAFO>
<PARAGRAFO etichetta="Cronologia">
  <CID etichetta="Data della ripresa">1930/ ca.</CID>
  <DTZG etichetta="Secolo">XX</DTZG>
  <DTSI_DTSV etichetta="Da" dsal="1929" dsuv="post">1929</DTSI_DTSV>
  <DTSE_DTSB etichetta="Da" dsal="1930" dsuv="post">1930</DTSE_DTSB>
  <DTMI etichetta="Motivazione">analisi storica/ analisi tecnico-formale</DTMI>
</PARAGRAFO>
<PARAGRAFO etichetta="Erovenienza">
  <FRVS etichetta="Stato">Italia</FRVS>
  <FRVP etichetta="Regione e provincia">Firenze</FRVP>
  <FRVC etichetta="Città">Firenze</FRVC>
  <FRCN etichetta="Raccolta">Collezione privata Sandberg Vavala Evelyn</FRCN>
  <EOL etichetta=""/>
</PARAGRAFO>
```

Figure 3. A portion of the XML output of a Scheda F

We are now looking into the available implementations of triple stores (probably choosing between OpenLink Virtuoso²⁴ and OwlIm by Ontotext²⁵) and a public SPARQL endpoint will be implemented next.

4.3 Linking triple to LOD

In order to make data non only open but also linked to the rest of the web, semi-automatic methods for link discovery and entity linking will be exploited (e.g. [24]).

The first task is to define datasets to be involved in the process (e.g. Dbpedia and Europeana, but we are also considering the Library of Congress – Linked Data Service and VIAF for the authorities, and also a few Italian resources such as CulturalItalia²⁶) and move towards a mechanism for ontology matching in order to establish relationships by way of alignment methods (see <http://www.ontologymatching.org/>).

In order to work in a real LOD dimension we plan to explore both the connection to other CH datasets and the integration of existent CH datasets. This means that new URLs of our entities and classes are being created and some others will be re-used from existing resources.

4.4 Create the browsing interface

The possibilities to explore the triple store are many. Each public dataset web-based uses different way to browse the data collection. Our idea is to create a responsive Web site based on a public faceted user interface²⁷.

The main purpose will be to make the interface as usable as possible, and to provide data exploration facilities through Semantic Web browsers²⁸.

5. MODELING THE SCHEDA F

The formalization of the schema of the Scheda F with a language suitable for exposing its data on the Web is going through an

accurate analysis of the domain we want to model and adopts appropriate ontology development processes, in order to guarantee the correct formalization in a short sequence of steps. Adopting a step-by-step formalization process is fundamental in our scenario, since proceeded incrementally in the development of the ontology, starting from modeling the most important concepts, namely, the Scheda F itself, the photograph, and the subject portrayed.

Since our final goal is to make available Scheda F data in a triple store of RDF statements, the target language we chose is OWL 2 DL [18]. In addition, the ontology development process we adopted is *SAMOD (Simplified Agile Methodology for Ontology Development)*²⁹, which involves ontology engineers and domain experts in each iteration of the process.

The material related to the development of the *F Entry Ontology (FEO)* (containing motivating scenarios, competency questions, glossary of terms, partial ontologies, SPARQL queries, examples, etc.) is already available online³⁰. The current version of the ontology is also available online³¹ and is further described in the following sections. An extension of the current version that adopts more CIDOC-CRM entities is planned, taking in consideration the analysis discussed in section 4.2.

5.1 The ontology

The current version of FEO introduces the classes and properties that characterize three specific concepts: the photograph, the work of art that is the subject of the photograph, and the Scheda F itself describing the photograph and its subject. During the first analysis, we realized a motivating scenario with examples accompanied by informal competency questions [21] and a glossary of terms [22].

The Scheda F is a document containing metadata describing a photograph having as subject a concrete work of art (or a part of it) or a group of different works of art. The collection has only one copy of each Scheda F, but there are multiple versions of the same photograph and multiple different photographs of the same work of art, which in turn may have undergone different modifications and changes captured by different photographs.

In order to capture correctly the complex maze of relationships and time- and context-dependent characterizations of the entity described, we describe Scheda F, photograph and work of art in the terms of the Functional Requirements for Bibliographic Records (FRBR), as follows:

- The Scheda F contains data about the cataloguing itself, which we are interested in, but is contained in a single version and we are not interested in its format. As such, we are only interested in its representation as FRBR Work and record its location as a FRBR Item.
- The photograph records a specific photographic event in a specific moment in time, but can be present in different forms in the collection (e.g., positive/negative, or different types of print, etc.), and of each form we may have multiple copies, each with its own story. As such, we represent the essence of every photograph as a FRBR Work, each form of such photograph as a FRBR Manifestation, and each individual copy as a FRBR Item.
- The work of art is a physical object that may have undergone different transformation events (deterioration, restoration, etc.). As such we represent the essence of each

²⁴ <http://virtuoso.openlinksw.com>

²⁵ <http://www.ontotext.com/owlim>

²⁶ A list in: <http://linkeddata.org/data-sets>

²⁷ As the OpenLink Virtuoso model used by Dbpedia.org: <http://dbpedia.org/fct/>

²⁸ A list in: <http://wiki.dbpedia.org/OnlineAccess#h28-13>

²⁹ <http://www.essepuntato.it/samod>

³⁰ <http://www.essepuntato.it/2014/03/fentry/samod>.

³¹ <http://www.essepuntato.it/2014/03/fentry>

work of art as a FRBR Work, the result of each transformation event as a FRBR Manifestation, and its physical characteristics as a FRBR Item.

All these characterizations are subject to authorial attribution as result of a production activity involving some agent in a specific moment in time.

In order to model these aspects, we imported a few existing ontologies:

- *FaBiO*, the *FRBR-aligned Bibliographic Ontology*³² [19], is an ontology based on FRBR originally developed for recording descriptions of bibliographic entities, both published and potentially publishable, or of entities that contain or are referred to by bibliographic references, or entities used to define such bibliographic references. FaBiO entities are primarily textual publications, such as books, magazines, newspapers, and journals, and items of their content, such as articles, conference papers and editorials. However, they also describe blogs, web pages, datasets, computer algorithms, images, metadata documents, experimental protocols, formal specifications and vocabularies, legal records, governmental papers, technical and commercial reports and similar publications, and also anthologies, catalogues and other collections;
- *Prov-O*, the *Provenance Ontology*³³ [20], that provides a set of classes, properties, and restrictions that can be used to represent provenance information about the activities (e.g., the creation of a photograph), the agents involved (e.g., the author of the photograph, the seller and the buyer in a transaction), and the entities they produce (e.g., the painting, the Scheda F, the photograph).

Through the conceptualization of the Scheda F, the intent has been to answer questions about relations between the three main concepts, according to the following informal competency questions [21]:

1. What are the *Schede F* and the photographs they describe?
2. What are the subjects portrayed in the photographs?
3. What are the entities responsible for the described objects, and what are their roles?
4. Which are the photographs that portray individual details of a larger work?
5. Which are the photographs owned by certain object in a specific time interval (e.g., the Federico Zeri Foundation or the personal collections from which they were drawn, etc.)?
6. When were the objects created and by whom (if any)?
7. Which are the characteristics of the location of the photographs, i.e., which photographs are included in a certain box (e.g., box 0014) and in which folder are they placed, and how are the folder organized in their box?

According to the general scenario introduced at the beginning of this section and its related competency questions, we have developed the version of FEO described in the following section.

5.1.1 Anatomy of FEO

The overall description of FEO is described in Figure 4, which shows a Graffoo diagram that describes the main classes and properties.

Graffoo, the *Graphical Framework For OWL Ontologies*³⁴, is an open source tool that presents graphically the classes, properties and restrictions of OWL ontologies as clear and easy-to-understand diagrams. In Graffoo:

- yellow rectangles identify classes;
- green rhomboids identify datatypes;
- pink circles identify individuals;
- black arrows identify assertions;
- blue arrows starting with a solid circle and having a solid arrow identify data properties; and
- green arrows starting with an empty circle and having an empty arrow identify object properties.

Arrows have solid lines if both domain and range are declared as the entities shown in the diagram, but they have dotted lines if the domain or the range are defined as other classes which the entities shown derive from. For instance, according to the diagram, the property *fentry:describes* can be used to link instances of the class *fentry:Photograph* to *fabio:AnalogueItem*, but its range actually is a more complex restriction of several classes.

FEO uses several entities defined in external ontologies, i.e., the PROV-O [20] (prefix *prov*), the OWL 2 DL version of FRBR³⁵ (prefix *frbr*), and FaBiO [19] (prefix *fabio*).

In order to show how to use the various entities of the ontology to describe vagueness/non-vagueness annotations, let us introduce the following natural language scenario:

A Scheda F shows that the Federico Zeri Foundation owns a photograph, authored by the publisher Bruckmann, portraying the Madonna and Child made by Leonardo da Vinci between 1470 and 1478. The photograph is part of the box 0360, issue 1.

5.1.2 The Scheda F

One of the main concepts of the ontology is the Scheda F. It is a particular metadata document containing metadata about a photograph and about the concrete objects portrayed by the photograph. It is defined in terms of FRBR Work (i.e., the class *fentry:FEntry*, a subclass of *fabio:EntityMetadata*) and of FRBR Item (as a computer file available online, i.e., the class *fabio:ComputerFile*). The object property *fentry:describes* links instances of the class *fentry:FEntry* to instances of any class (belonging to any FRBR layer), among which instances of *fentry:Photograph*, *fabio:ArtisticWork*, etc.

They are defined as follows (in Manchester Syntax):

```
Class: fentry:FEntry
  SubClassOf: fabio:EntityMetadata

ObjectProperty: fentry:describes
  Domain: fentry:FEntry
  Range:
    fabio:Work or fabio:Expression or
    fabio:Manifestation or fabio:Item
```

In our example, the Scheda F contains information about both the photograph and the work of art, and therefore its link to the photograph metadata and the related portrayed object is defined as follows (in Turtle):

```
Individual: f-entry-1
  Types: fentry:FEntry
  Facts:
```

³² <http://purl.org/spar/fabio>

³³ <http://www.w3.org/TR/prov-o/>

³⁴ <http://www.essepuntato.it/graffoo>

³⁵ <http://purl.org/spar/frbr>

```
fentry:describes madonna-and-child-item ,
fentry:describes madonna-and-child-photo-item ,
fentry:describes madonna-and-child-photo-work ,
fentry:describes madonna-and-child-work
```

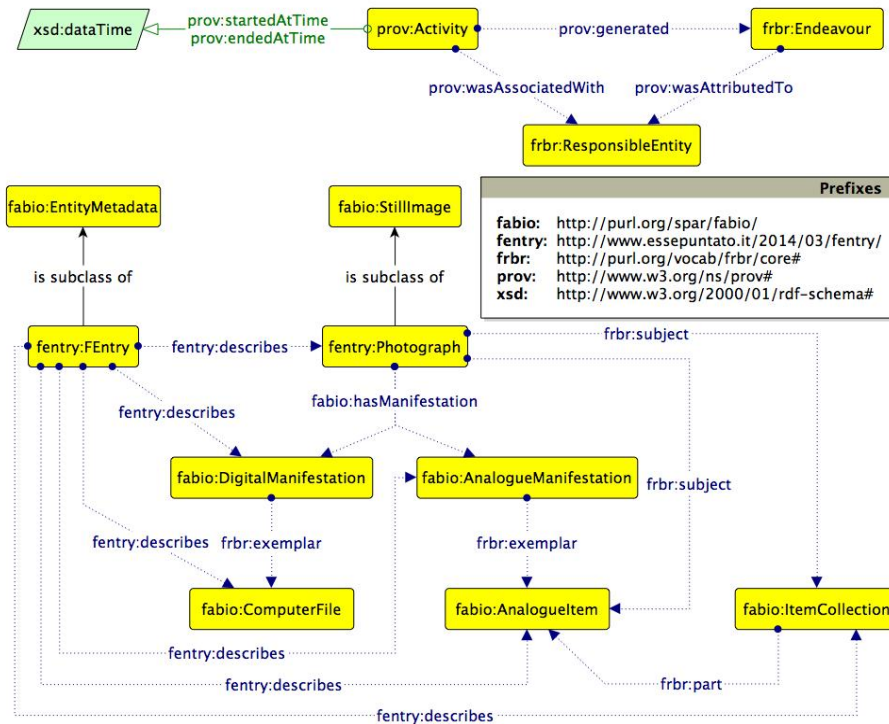


Figure 4. The Graffoo diagram (<http://www.essepuntato.it/graffoo>) of the main classes and relations defined in FEO

5.1.3 The photograph

The photograph is defined as a static, visual representation of some concrete object (or of a part of it or of a group of different objects). It is defined in terms of FRBR Work (i.e., the class *fentry:Photograph*, a subclass of *fabio:StillImage*), FRBR Manifestation (i.e., the classes *fabio:DigitalManifestation* and *fabio:AnalogueManifestation*), and FRBR Item (i.e., the classes *fabio:ComputerFile* and *fabio:DigitalItem*). The FRBR property *frbr:subject* links the photograph to the concrete object it portrays. The class *fentry:Photograph* is defined as follows:

```
Class: fentry:Photograph
SubClassOf:
  fabio:hasManifestation only (
    (fabio:AnalogManifestation or
     (fabio:DigitalManifestation and
      (frbr:exemplar only fabio:ComputerFile))),
  fabio:StillImage,
  frbr:subject some (
    (fabio:AnalogItem or frbr:Object or frbr:Place or
     (fabio:ItemCollection and
      (frbr:part only fabio:AnalogItem))))
```

In our example, the photograph metadata, the actors involved (i.e., the owner of the photograph and its author, by means of the property *prov:wasAttributedTo*), and the physical containment of the photo in terms of issues and boxes (by means of the property *frbr:partOf*), are described as follows:

```
Individual: madonna-and-child-photo-item
Types: fentry:AnalogItem
Facts:
  frbr:partOf box0360-issue1 ,
  prov:wasAttributedTo federico-zeri-foundation
```

```
Individual: madonna-and-child-photo-work
Types: fentry:Photograph
Facts:
  prov:wasAttributedTo bruckmann ,
  frbr:subject madonna-and-child-item ,
  fabio:hasPortrayal madonna-and-child-photo-item
```

```
Individual: box0360
Types: fentry:AnalogItem
```

```
Individual: box0360-issue1
Types: fentry:AnalogItem
Facts: frbr:partOf :box0360
```

```
Individual: bruckmann
Types: frbr:ResponsibleEntity
```

```
Individual federico-zeri-foundation
Types: frbr:ResponsibleEntity
```

5.1.4 Portrayed work of art

The concrete object portrayed by a photograph is described in terms of the FRBR description provided by FaBiO. In particular, it is defined in terms of FRBR Work (i.e., the class *fabio:ArtisticWork*), FRBR Manifestation (i.e., the class *fabio:AnalogueManifestation*), and FRBR item (i.e., the class *fabio:AnalogItem*).

In our example, the object portrayed by the photograph, its author and the creation process that resulted in the generation of such object, are described as follows:

```
Individual: leonardo-da-vinci
Types: frbr:ResponsibleEntity
```

```

Individual: madonna-and-child-creation
Types: prov:Activity
Facts:
  prov:startedAtTime
    "1470-01-01T00:00:00"^^xsd:dateTime ,
  prov:endedAtTime
    "1478-01-01T00:00:00"^^xsd:dateTime ,
  prov:wasAssociatedWith leonardo-da-vinci ,
  prov:generated :madonna-and-child-work

```

```

Individual: madonna-and-child-item
Types: fentry:AnalogueItem

```

```

Individual: madonna-and-child-work
Types: fabio:ArtisticWork
Facts:
  prov:wasAttributedTo leonardo-da-vinci ,
  fabio:hasPortrayal madonna-and-child-item

```

5.1.5 Using the imported ontologies

Using the aforementioned entities, we provide a description of any Scheda F of the Federico Zeri Foundation, at least at a general level. The use of FaBiO and PROV-O allows us to specify additional entities such as types of activities (e.g., *prov:Copyright*, *prov:Create*, *prov:Publish*), types of responsible entities (e.g., *frbr:Person*, *frbr:CorporateBody*), and types of works (e.g., *fabio:ResearchPaper*, *fabio:WorkCollection*), that we do not show here because of space limits.

Of course, additional analysis may be required so as to extend FEO to use all the CIDOC-CRM entities relevant to all the fields that the schema of the Scheda F makes available, and in order to use the Publishing Roles Ontology (PRO)³⁶ [23] that describes the roles that responsible entities have in relation to an object.

While the extension of FEO will result in more CIDOC-CRM entities to be involved, the first results of our capillary investigation is introduced in the following section.

5.2 The mapping ICCD / CIDOC-CRM

A deep analysis of the schema of the Scheda F show that it is organized in semantically independent sections (called “paragrafi”, or *paragraphs*), each one belonging to a specific FRBR concept (Work, Manifestation, Item); this allows the mapping to proceed by logical sections affecting only a limited number of entities and relating these entities to the data documented by the fields of the schema (see a fragment of full mapping in Figure 5).

A deep and complete technical description of the mapping details are beyond the scope of this paper, which aims to give a general presentation of the project; thus the following sections give a summary view of the conversion and of the approach applied in translating the schema to CIDOC-CRM.

5.2.1 Photo objects and Scheda F

The schema of the Scheda F concerns the description of photographic pictures which are, from the CIDOC-CRM perspective, physical objects purposely created by human activity, so the class used to represent it is of course *E22 Man-Made Object* (we will refer to it as the photograph). For storing the meta-properties about the schema itself (e.g. creator, date, schema type, etc.) we chose the class *E31 Document* which documents (by property *P70 documents*) the photography object, referring to it as the entry.

Since the photos in the Zeri Foundation photographic catalogue are all about artworks, the schema of the Scheda F contains links to the schema of the Scheda OA which describes the work of art itself; since the schema of the Scheda OA is not in scope of this paper, we will not cover it here in details. Suffice to say that the Scheda OA is classed, similarly to the Scheda F, as an *E31 Document* and the entry links to it by property *P67 refers to*.

5.2.2 Work level

The work of art is the subject of the photography and, from the CIDOC-CRM perspective, it is a physical object purposely created by human activity; as such it is classed as an *E22 Man-Made Object* as well and, and it is linked to the photo by property *P62 is depicted by*.

The author of the photo object can either be a person or a corporate body so we can not deepen into CIDOC-CRM hierarchy beyond *E39 Actor*, which is then linked to the peculiar activity it belongs to: e.g. the main author is linked by *P14 performed to E12 Production* activity, which is then linked to the photo by property *P108 has produced*; alternative authors can be linked to the photo via different activities.

5.2.3 Manifestation level

The description of the object is, from an FRBR perspective, at the Manifestation level, documenting specific properties of the object itself, such as:

- *Material*: materials of which the object is made, mapped through property *P45 consist of to E57 Material*
- *Dimensions*: height, width, thickness, diameter, mapped to actual values as *E54 Dimension* through properties *P43 has dimension*, *P90 has value*, and *P91 has unit*; for example: photo object (*E22*) *has dimension (P43)* height of photo (*E54*) *has unit (P91)* mm (*E58*), *has value (P90)* 227 (*E60*).
- *Technique*: the technique used in producing the photo object is represented by a link to *E12 Production* (through property *P108 was produced by*) which is then linked via *P32 used technique* to the controlled ICCD vocabulary mapped in *E55 Type*.

Beside the physical object data, the schema of the Scheda F allows the cataloguer to describe the photo creation by means of the date and place of the shot. This is represented by a link through *P94 was created by* to the activity *E65 Creation* and *P108 was produced by* to *E12 Production*, which are then linked to *E53 Place* by *P7 took place at* and to *E4 Period* by *P10 falls within*.

Other data belonging to Manifestation level registered by the schema of the Scheda F are the type of the object (e.g. positive, negative, etc.) which is naturally mapped to the ICCD controlled vocabulary on an *E55 Type* by property *P2 has type*; the number of physical items composing the object itself which is represented by *E60 Number* linked through *P57 has number of parts*; the possible relations with other photo objects (de-composition for complex or related photos), though logically are a *P46 is composed of (forms part of)* to other standalone objects, are actually represented in the schema as a numeric suffix to the identifier, so it has been represented by a *E42 Identifier* (through *P1 is identified by*) which can have a *P3 has note* to *E62 String*.

Copyright data also belong to the Manifestation level and are linked through *P104 is subject to* to a *E30 Right* which *P75 is possessed by* a *E39 Actor*. That link is possible being *E22 Man-Made Object* a descendant both of *E71 Man-Made Thing* and *E72 Legal Object*.

³⁶ <http://purl.org/spar/pro>

5.2.4 Item level

The schema of the Scheda F includes a few fields that document the current location (both geographical and logical) of the photo object and additional former locations if known. These locations are, in CIDOC-CRM, classed as *E53 Place*, even the logical ones (e.g., the shelfmarks). We chose *P54 has current permanent location* for linking the current location to the photo, and *P53 has former or current location* to describe the former locations. Nested locations (e.g. the shelfmark is contained into a geographically positioned building, or a logical folder is inside a logical box) are linked to each other with property *P89 falls within*.

Another important information belonging to the Item level is the physical conditions and state of preservation of the object. To map these we chose the class *E3 Condition State* and we linked it to the photo by property *P44 has condition*. Since the condition states in the schema come from a closed vocabulary, *E55 Type* has been used to map it for populating the field.

5.2.5 Properties of the Scheda F

The properties of the Scheda F consist mainly of administrative data, the most important being the unique identifier within the catalogue, which is mapped to *E42 Identifier* through property *P48 has preferred identifier*.

The *entry type* identifies the type of the schema among the ones provided by ICCD (so in this case the value is always “F”) and is naturally mapped to a *E55 Type* via property *P2 has type*.

The corporate group which is responsible for safeguarding the photo object is mapped to *E40 Legal Body* via property *P50 has current keeper*.

Other administrative data describe the individual cataloguers who operated on the entry and the chronological history of the edits, such as, for example, the first draft responsibility: this action is logically mapped to *E65 Creation* class through *P94 was created by* property, and the action itself linked by *P4 has time-span* to *E52 Time-Span* and by *P14 carried out by* to *E39 Actor*, which store respectively the temporal coordinates and the responsible actor name.

CD - Codici	Livello catalogazione	LIR	
	Codice univoco: Codice regione	NCTR	
	Codice univoco: Numero catalogo generale	NCTN	P48 has preferred identifier - E42 Identifier
	Ente schedatore	ESC	
OG - Bene culturale	Ente competente per tutela	ECP	P50 has current keeper - E40 Legal Body
	Definizione bene: Definizione	OGTD	P2 has type - E55 Type
RV - Relazioni	Quantità: Quantità degli esemplari	QNTN	P57 has number of parts - E60 Number
	Struttura complessa: Livello nella struttura gerarchica	RVEL	P1 is identified by (identifies) - E42 Identifier
LC - Localizzazione geografico-amministrativa	Struttura complessa: Note	RVEZ	P1 is identified by (identifies) - E42 Identifier P3 has note - E62 String
	Localizzazione: Regione	PVCR	P54 has current permanent location (is current permanent location of) - E53 Place P87 is identified by - E48 Place Name
	Localizzazione: Provincia	PVCP	P54 has current permanent location (is current permanent location of) - E53 Place P87 is identified by - E48 Place Name
	Localizzazione: Comune	PVCC	P54 has current permanent location (is current permanent location of) - E53 Place P87 is identified by - E48 Place Name
	Collocazione specifica: Denominazione contenitore fisico	LDCN	P54 has current permanent location (is current permanent location of) - E53 Place P87 is identified by - E48 Place Name
	Collocazione specifica: Indicazioni viabilistiche	LDCU	P54 has current permanent location (is current permanent location of) - E53 Place P87 is identified by - E45 Address
	Collocazione specifica: Denominazione contenitore giuridico	LDCM	P54 has current permanent location (is current permanent location of) - E53 Place P87 is identified by - E46 Section Definition
	Collocazione specifica: Informazioni specifiche sulla collocazione	LDCS	P54 has current permanent location (is current permanent location of) - E53 Place P3 has note - E62 String
	Tipo di localizzazione	TLC	P53 has former or current location (is former or current location of) - E53 Place P2 has type - E55 Type
	Altra localizzazione: Stato	PRVS	P53 has former or current location (is former or current location of) - E53 Place P87 is identified by - E48 Place Name
	Altra localizzazione: Provincia	PRVP	P53 has former or current location (is former or current location of) - E53 Place

Figure 5. A fragment of the full mapping between the schema of the Scheda F and CIDOC-CRM

6. FUTURE WORKS AND CONCLUSIONS

In this paper we presented the first steps of the *Zeri e LODE* project towards a comprehensive plan to make available in Linked Open Data the information coming from one of the most important collections of photographs of works of art in the world, the Photo Archive of the Federico Zeri Foundation.

In this paper we especially focused on the most innovative aspect of the project, the development of a mapping between the conceptual model of the database describing the collection, which is based on the Scheda F of the *Istituto Centrale per il Catalogo e la Documentazione* (ICCD), onto CIDOC-CRM, and of the interpretation issues we found. Subsequent work will include the creation of the triple store, the creation of the conversion tools from the existing database to the triple store, the generation of the links from the data contained in the triple store to the information available on the net (in particular to DBPedia and other authority files), and the generation of a navigable and queryable interface to the triple store.

The idea of the project is to develop a model that could be exploited in other different collections in the photographic domain.

So, although in our mind the most important building blocks for the *Zeri e LODE* project have already been defined and set, there is still plenty of work to do. Additional details in the completed work may end up requiring some serious reworking. For instance, the issue of restricting allowable values in fields that are open in the Schema of the Scheda F for ICCD, but use a controlled vocabulary in the implementation of the Zeri Foundation will have to be understood and added to our implementation of the FEO ontology, and similarly the foreseen modifications and upgrades of the Scheda F will have to be understood and integrated in the FEO as well. Finally, our current use of the CIDOC-CRM entities is somewhat limited and constrained, and additional mappings are foreseen and looked forward to. And yet, we still have not answered to the most pressing question of all: are the entities in CIDOC-CRM even sufficient for our needs? Shall we need to invent new ones and actually extend CIDOC-CRM?

7. REFERENCES

- [1] N. Crofts *et al.* (2011, Nov.). *Definition of the CIDOC Conceptual Reference Model (version 5.0.4)* [Online]. Available: http://www.cidoc-crm.org/docs/cidoc_crm_version_5.0.4.pdf
- [2] M. Damova *et al.*, “Creation and Integration of Reference Ontologies for Efficient LOD Management,” in *Semi-Automatic Ontology Development: Processes and Resources*. Hershey PA: IGI Global, 2011, pp. 162-199. DOI= 10.4018/978-1-4666-0188-8.ch007
- [3] V. De Boer *et al.*, “Supporting Linked Data Production for Cultural Heritage Institutes: The Amsterdam Museum Case Study,” in *The Semantic Web: Research and Applications* (LNCS, vol. 7295). Berlin Heidelberg, Germany: Springer-Verlag, 2012, pp. 733-747.
- [4] M. Doerr. (2010, Dec. 10). *The Europeana Data Model (EDM) mapping to CIDOC-CRM* [Online]. Available: http://www.cidoc-crm.org/docs/CRM-EDM_FRBR.ppt
- [5] J. Edelstein *et al.* (2013). *Linked Open Data for Cultural Heritage: Evolution of an Information Technology* [Online]. Available: <http://www.whysel.com/papers/LIS670-Linked-Open-Data-for-Cultural-Heritage.pdf>
- [6] *EDM (Europeana Data Model) documentation* [Online]. Available: <http://pro.europeana.eu/edm-documentation>
- [7] M. Fernández-López *et al.*, “Methontology: From Ontological Art Towards Ontological Engineering,” in *Proc. Symp. Ontological Engineering AAAI*. Menlo Park, California: AAAI Press. AAAI Technical Report SS-97-06, 1997.
- [8] *Fondazione Federico Zeri – Fototeca* [Online]. Available: http://www.fondazionezeri.unibo.it/home/fototeca/00000045_la_fototeca.html
- [9] *FRBR (Functional Requirements for Bibliographic Records)* [Online]. Available: <http://www.ifla.org/publications/functional-requirements-for-bibliographic-records>
- [10] *FRBR-CRM, FRBRoo* [Online]. Available: http://www.cidoc-crm.org/frbr_inro.html
- [11] ICCD (Istituto Centrale per il Catalogo e la Documentazione), *Beni fotografici* [Online]. Available: <http://www.iccd.beniculturali.it/index.php?it/387/beni-fotografici>
- [12] A. Jentzsch *et al.* (2011, Sept. 9). *State of the LOD Cloud (version 0.3)* [Online]. Available: <http://lod-cloud.net/state/>
- [13] P. Le Boeuf *et al.* (2012). *Definition of the CIDOC Conceptual Reference Model* [Online]. Available: http://www.cidoc-crm.org/official_release_cidoc.html
- [14] T. Lebo *et al.* (2013, April 30). *PROV-O: The PROV Ontology. W3C Recommendation* [Online]. Available: <http://www.w3.org/TR/prov-o/>
- [15] J. Marden *et al.*, “Linked open data for cultural heritage: evolution of an information technology,” in *Proc. 31st ACM Int. Conf. Design communication*. New York, NY: ACM, 2013, pp. 107-112.
- [16] B. Motik *et al.* (2012, December 11). *OWL 2 Web Ontology Language: Structural Specification and Functional-Style Syntax (Second Edition)*. W3C Recommendation [Online]. Available: <http://www.w3.org/TR/owl2-syntax/>
- [17] *Pan-Canadian Documentary Heritage Network (PCDHN) Linked Open Data (LOD) Visualization “Proof-of-Concept”* [Online]. Available: http://www.canadiana.ca/sites/pub.canadiana.ca/files/PCDHN%20Proof-of-concept_Final-Report-ENG_0.pdf
- [18] S. Peroni *et al.*, “Reflecting on the Europeana Data Model,” in *Digital Libraries and Archives* (CCIS, vol. 354). Berlin Heidelberg, Germany: Springer-Verlag, 2012, pp. 228-240.
- [19] S. Peroni *et al.*, “The aggregation of heterogeneous metadata in Web-based cultural heritage collections. A case study,” in *International Journal of Web Engineering and Technology*, vol. 8, n. 4, pp. 412-432, Feb. 2013. DOI=10.1504/IJWET.2013.059107
- [20] S. Peroni and D. Shotton, “FaBiO and CiTO: Ontologies for describing bibliographic resources and citations,” in *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 17, pp. 33-43, Dec. 2012.
- [21] S. Peroni *et al.*, “Scholarly publishing and linked data: describing roles, statuses, temporal and contextual extents,” in *Proc. 8th Int. Conf. Semantic Systems*, New York, NY: ACM, 2012, pp. 9-16.
- [22] T. Scarselli *et al.* (2012). *Mapping ICCD Archaeological Data to CIDOC-CRM: the RA Schema* [Online] Available: http://www.academia.edu/5549957/Mapping_ICCD_archaeological_data_to_CIDOC-CRM_the_RA_Schema
- [23] M. Uschold and M. Gruninger, “Ontologies: principles, methods and applications,” *Knowledge Engineering Review*, vol. 11, pp. 93-136, 1996.
- [24] J. Volz *et al.*, “Discovering and maintaining links on the web of data,” in *Proc. 8th Int. Semantic Web Conf.*, Berlin Heidelberg, Germany: Springer-Verlag, 2009, pp. 650-665.