Reflecting on the Europeana Data Model

Silvio Peroni[1], Francesca Tomasi[2], and Fabio Vitali[1]

[1] Department of Computer Science, University of Bologna, Italy
{essepuntato,fabio}@cs.unibo.it
[2] Department of Classical Philology and Italian Studies, University of Bologna, Italy
francesca.tomasi@unibo.it

Abstract. We describe some issues arising while using Europeana, and analyze some features of the Europeana Data Model (EDM), starting from the rationale of the project. Some aspects of the theoretical model, derived mostly from the mapping between the provided Cultural Heritage Object (CHO) and the EDM, prevent useful results in users' queries. The concept of media type, the multi-layer description and the relation between roles and values are some issues about which we reflected. The aim of Europeana to make records available as Linked Open Data on the Web could require moreover a redefinition of the implementation techniques.

Keywords: EDM, Linked Data, RDF, DC, CIDOC CRM, FRBR.

1 Introduction

Europeana[1] is the European Digital Library, a distributed access point to Europe's multilingual cultural heritage in a digital form. The main aim of the project is to collect metadata from a large number of providers, mainly cultural institutions, across Europe, and to enable search and discovery of cultural items described therein.

The metadata aggregation is based on a mapping between the providers' data description and the Europeana model. The Europeana v1.0 project[2] [1] proposes the Europeana Data Model (EDM)[3] that defines a set of classes and properties to be used in Europeana for describing cultural objects. The EDM [2] is a clear improvement over the earlier data model, the Europeana Semantic Elements (ESE) [3]. ESE was meant to express the providers' datasets using the Dublin Core (DC) standard as its "lowest common denominator", while EDM, based on the DCMI Metadata Terms and a number of more advanced metadata models, adopts "an open, cross-domain Semantic Web-based framework" leaving each provider free to use their preferred metadata standard with regard to the element sets and the vocabularies of values [4].

[1] User access: http://europeana.eu/portal
[2] http://pro.europeana.eu/web/europeana-v1.0.
In: Europeana Professional: http://pro.europeana.eu
[3] The family of technical documents about EDM (in particular Definition, Primer, Guidelines) could be found at:
http://pro.europeana.eu/web/guest/edm-documentation

M. Agosti et al. (Eds.): IRCDL 2012, CCIS 354, pp. 228–240, 2013.

Given this, an extreme heterogeneity can be observed in the descriptions of cultural objects, a situation determined by the differences in existing collections that Europeana involves (museums, archives, audiovisual collections and libraries), and by the different kinds of objects described in the cultural repositories that were harvested (i.e. manuscripts, documents, paintings, art and architecture objects, photos, videos, etc.). Also, the variety of descriptive situations depends both on the reference models for the metadata element sets (not only DC, SKOS, and the CIDOC CRM, but also e.g., EAD for archives, METS for digital libraries, TEI for literary texts[4]) and from the vocabularies of values (authority lists, thesauri or controlled vocabularies) used by data providers (such as Geonames, Art&Architecture Thesaurus, Iconclass, WordNet, Dewey Decimal Classification, DBPedia, etc.)[5].

Furthermore, even if the EDM has been introduced, a majority of records in Europeana still seems to follow the old ESE data model and the users' query interface is based on just a few categories of Dublin Core. The main problems in using the current release of Europeana derives in part from the above-mentioned issues: the original object descriptions are often lost and the ESE, the first proposed model, was not sufficient to describe the complexity of many cultural objects, as it was based on just a few DC categories. Providers that had complex and structured descriptions have had to force them into to a much simpler model and providers that created descriptions in a model not compliant with ESE have lost data or have forced them into incorrect properties, leading to aggregations of imprecise information. Consequently, many user queries cannot be satisfied completely. Probably the complete integration of EDM in Europeana records will help towards completeness and correctness of the contained information. But some other improvements could solve many situations that not even EDM takes into complete account.

The whole point of this effort clearly is to allow users to enact better queries and to obtain better results through the use of a sophisticated and increasingly ontological metadata model on which to let the search engine work. Europeana is now working on improving the quality of the responses to users' queries, but much is still to be done.

For example at the moment there is no possibility to perform a multilingual search, although the vocabulary alignment is a problem being studied currently by the Europeana group, since one of the aims of subproject EuropeanaConnect[6] is in fact to solve this gap.

---

[4] DC, SKOS and CIDOC CRM will be discussed in section 2. As regards to the other schemas we just mention here the official sites: EAD (Encoded Archival Description), `http://www.loc.gov/ead/`; METS (Metadata Encoding & Transmission Standard), `http://www.loc.gov/standards/mets/`; TEI (Text Encoding Initiative), `http://www.tei-c.org`

[5] A comprehensive list of metadata and vocabularies published ad Linked Data sets could be found in: W3C Incubator Group Report 25 October 2011. Library Linked Data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets at:
`http://www.w3.org/2005/Incubator/lld/`
`XGR-lld-vocabdataset- 20111025/`

[6] `http://www.europeanaconnect.eu/`

Some other features are planned to be implemented using Semantic Web technologies[7]. For example, Europeana lacks a semantic network for the subjects, that could help users in finding records by specifying either the exact word or any of its synonyms, hyponyms, hypernyms, related terms, etc. WordNet, for example, has been published as a RDF vocabulary[8] and could be used after a vocabulary alignment. Another issue that Semantic Web technologies could solve is to enrich the existing metadata sets, contextualising objects and using authoritative sources, including controlled vocabularies, concretely shared in an integrated environment. The "related content", presented in each object description in Europeana, is now mostly limited to other objects from the same data provider. More complex and structured relationships between what Europeana calls the provided "Cultural Heritage Object (CHO)" [2] and other pertinent resources, internal (other data providers) or external (other digital libraries), could be solved in a Linked Data perspective.

Although some features are now being studied as an experimental increment to the existing feature set, some discrepancies could be noticed between the EDM and the objects described, for which we do not know of any on-going work.

In this paper we analyze the query results of the current implementation and propose some reflections on the EDM in this phase of implementation. We focus here on three aspects, derived from the adoption of DC as the end-user property set: the "media type" concept, the multi-layer description of subjects and the connection between roles and values (Section 3). Even if the original adoption of DC is the main reason for the limitations in user queries, a redefinition of some classes and properties of EDM could solve some issues. Finally, given that one of the main aims of Europeana project is to expose metadata as Linked Open Data we verified that the current implementation is fairly good for the data that has been converted, but is still lacking in handling properly error cases. For this reason we analysed the "data.europeana.eu" Linked Data responses and the issues with empty resources (Section 4).

2 Related Works

In its specifications (EDM), Europeana mentions vocabularies, models and ontologies adopted in its data model [2]. The aim is to represent metadata for cultural heritage objects and to give access to digital representations of these objects. The EDM moves in the context of data aggregation, where objects can be complex, and several data providers may entertain different views on them.

The basis of the metadata description is RDF statements. An XML Schema has been defined for describing classes and properties. Some classes and properties are reused from public models: DC, DCterms, SKOS, OAI-ORE, CIDOC-CRM, FRBR. Some other classes and properties are specifically created for the EDM and are mostly equivalent to predicates used in the most common ontologies.

[7] Library Linked Data Incubator Group wiki. Use Case Europeana: `http://www.w3.org/2005/Incubator/lld/wiki/Use_Case_Europeana`

[8] Wordnet 3.0 in RDF: `http://semanticweb.cs.vu.nl/lod/wn30/`

In addition to classes and properties, Europeana is defining also controlled vocabularies useful for CHO interoperability (such as AAT, DDC, DBpedia, Iconclass). The main aim of Europeana is to work on Linked Data both exposing record sets [5] and using Linked Data resources [6] in order to augment Europeana content.

In the following sub-sections we introduce the main external models adopted by EDM, highlighting which part of them are effectively used.

2.1 Dublin Core

The current versions of the Dublin Core (DC) Metadata Elements [7] and of the DC Metadata Terms [8] are the most widely used vocabularies for describing and cataloguing resources. These vocabularies have become particularly important and relevant for sharing metadata about documents among different repositories and digital libraries. While very useful for creating basic metadata that permit bibliographic resource descriptions (e.g., *creator*, *contributor*, *publisher*, *format*), the main limitation of DC is a consequence of the generic nature of its terms. In fact, its classes are organised without a strong hierarchical structure and their properties often lack in clear domain/range definitions. EDM makes extensive use of DC Elements and DC Terms entities, such as the properties *dc:subject*, *dc:contributor*, *dcterms:created* and *dcterms:alternative*.

2.2 SKOS

Data providers, publishers and aggregators, such as Europeana, need to classify the resources they publish according to discipline-specific thesauri and classification schemes. The Simple Knowledge Organization System (SKOS) [9] is an RDFS ontology to support the use of knowledge organization systems (KOS). A large number of well-known thesauri and classification systems have started to convert their specifications into SKOS documents, such as the "Nuovo Soggettario" of the National Central Library in Florence[9]. This makes SKOS the de facto standard for encoding controlled vocabularies for the Semantic Web. EDM uses the main SKOS class, i.e. *skos:Concept*, defined as a particular kind of *edm:NonInformationResource* for introducing an idea or notion.

2.3 FRBR

The Functional Requirements for Bibliographic Record (FRBR) [10] is a general model for describing bibliographic entities, such as documents and artistic works. FRBR specifies four basic concepts – work, expression, manifestation and item – used for characterising a particular bibliographic entity from different perspectives. In particular:

- A *work* is the *abstract essence* of an intellectual or artistic creation, e.g. the ideas in Shakespeare's head concerning the *Macbeth*. A work is realised in one or more expressions;

[9] `http://thes.bncf.firenze.sbn.it/`

- An *expression* is the *content* of a particular work at a specific point in time, e.g. the final text of the *Macbeth* written by Shakespeare or its Italian translation made by Andrea Maffei. An expression is embodied in one or more manifestations;
- A *manifestation* is the particular *format* in which an expression is stored, such as a printed object or a digital document, e.g. the 2005 edition of *Macbeth* published by Penguin Books or its HTML Italian version published by. A manifestation is exemplified in one or more item;
- An *item* is a particular *physical or electronic copy* of the *Macbeth* that a person can own, e.g. the printed version of that book you have in your bookcase or the specific HTML document of its Italian version you are visualising in your browser.

Overall, EDM makes only limited use of FRBR concepts, although it declares explicitly their adoption. The only specific references to FRBR are:

- the class *edm:InformationResource*, defined as union of *FRBR Work*, *FRBR Expression* and *FRBR Manifestation* that results in collapsing completely the hierarchy of the FRBR model;
- the classes *edm:Event* and *edm:Place*, defined as equivalent to *FRBR Event* and *FRBR Place* respectively.

2.4 ORE

The Open Reuse and Exchange specification (ORE specification) [11] is a standard defined by the Open Archives Initiative for describing and exchanging aggregations of Web resources. Europeana uses two terms from this model:

- *Aggregation*, i.e. a particular resource that aggregates, either logically or physically, other resources;
- *Proxy*, used to refer to a specific aggregated resource in a context of a particular aggregation.

EDM uses all the main classes and properties of the ORE specification. For instance, it allows one to describe a "cultural heritage object" (i.e., *edm:providedCHO*) and its digital representations (i.e., *edm:WebResource*) as a particular aggregation (*ore:Aggregation*) representing the results of the activity of a particular data provider (i.e., *edm:Agent*).

2.5 CIDOC CRM

CIDOC Conceptual Reference Model (CRM) [12] is an ISO standard defining a model for describing and sharing cultural heritage information. It provides entity definitions and a formal multi-level structure to link physical objects to related events and

agents (i.e., people and organisations), so as to represent a mediator between different sources of cultural heritage information (e.g., museums, libraries and archives).

EDM aligns some of its classes and properties to the CIDOC CRM specification, for instance the class *edm:Event* as equivalent to *E4 Period*, the class *edm:InformationResource* as subclass of *E73 Information Object*, and the property *edm:wasPresentAt* as equivalent to *P12 occurred in the presence of (was present at)*[10].

3 Issues Arising While Using Europeana

The EDM rationale is based on some principles [14]:

1. distinction between "provided objects" (painting, book, movie, etc.) and their digital representations;
2. distinction between objects and metadata records describing an object;
3. allow for multiple records for a same object, containing potentially contradictory statements about it;
4. support for objects that are composed of other objects;
5. compatibility with different levels of description;
6. standard metadata format that can be specialized;
7. support for contextual resources, including concepts from controlled vocabularies.

Although the Europeana core classes stress the difference between the provided object (*edm:ProvidedCHO*), i.e, the "real object", and its digital representation (*edm:WebResource*), i.e., its Web resource, sometimes this difference is not evident at all in the aggregated metadata exposed to the final user, generating confusion. Sometimes the description seems to be addressed to the electronic version, some other to the original work, without a clear distinction (see 3.1). Additionally, the Europeana contextual classes, which are designed to answer to the four fundamental questions of the *who* (the Agent), the *where* (the Place), the *when* (the Time), and the *what* (the Concept) of the object, sometimes are not correctly represented in the metadata description because of a potential multi-layer representation issue derived from the stratification of object and subject (see 3.2). Therefore, the rationale of the EDM appears not always respected and the application of the listed properties is not totally functional (see 3.3). Here we describe some examples of these limits.

3.1 The Media Type

The first and most evident source of confusion is the concept of media type found as the topmost choice in the filter section after every query (*edm:type* = text, image, audio, video). The media type is sometimes congruent with the type of the provided

[10] While in the Europeana Data Model the CIDOC CRM property *was present at* has the identifier *P121*, in [12] that property is defined as *P12 occurred in presence of (was present at)*.

CHO, and sometimes to its web resource. Yet, the rationale of Europeana is to distinguish between the description of the CHO and its digital representation. If we search for any object called "illuminated manuscript" we receive different answers: sometimes it has type "text", sometimes type "image" some other times "physical object". In general, resources classified as IMAGE are in fact image files (regardless of whether they represent pictures or physical objects such as buildings or statues or manuscripts), but resources classified as TEXT are sometimes texts, and sometimes images of texts (e.g., photographs of old volumes or manuscripts). One may wonder which would be more useful for searches, i.e., for the media type to refer to the web resource, providing a description of a computer-specific object, or to the cultural heritage object, which is what the user would be actually searching for. In both cases, it would be quite important to provide subtypes: they could be either subtypes of the relevant Internet MIME type[11] in the first case, or a selection of the values found in the *dc:type* facet of the records as specified in the collections in the other case.

3.2 Multi-layer Descriptions

The issue of the separation between web resource and cultural heritage object can be subsumed in the issue of separating objects and subjects in record descriptions. In fact, what does exactly a Europeana record describe? Is it an image, the content of the image, or the object represented in the content of the image? Sometimes this is easy to understand, and sometimes it creates interpretation issues, and the problem of distinguishing between object and subject in a record can go several layers deep. For instance, consider the 1756 publication by Giambattista Piranesi called "Le antichità romane", containing prints of famous Roman monuments, including the Coliseum. A best seller of the time, the volume appears in several of the collections of Europeana. We analysed 3 items, alla 3 form the Bildarchiv Foto Marburg as data provider. According to one item[12], the page representing the Coliseum is of *dc:type* druck (print), *dc:creator* Giambattista Piranesi, *dc:date* 1756. According to another[13], the same page is of *dc:type* amphitheatre, *dc:description* Location:Rome and its *dc:date* is 70/80 a.D. (and no *dc:creator*), but reports (in the *dc:description* field) that the actual photo was taken in 1956, and that the content is an extract of the Piranesi's book of 1756. This is coherent with several colour photographs of the Coliseum[14] present in the same collection, whose *dc:type* is also amphitheatre, *dc:date* is 72/80 a.D., further adding that *dc:format* is travertine, and *dc:contributor* is Vespasianus (as contractor).

In cases such as Piranesi's, the number of layers of subjects is multiple: the CHO being described is a 1956 b/w photograph of unknown creator, whose subject is a

[11] http://www.iana.org/assignments/media-types/text/index.html
[12] http://www.europeana.eu/portal/record/08501/7B74073B6E9E90F5B572EF6DF20426AF0135202E.html
[13] http://www.europeana.eu/portal/record/08501/EA45A0B5F838ABDDD1956DE3BE636A70F1B8EA8A.html
[14] http://www.europeana.eu/portal/record/08501/43E4B1EF54983567EC92DDCDD57B3DBD2D4CC013.html

1756 print whose creator is Giambattista Piranesi, whose subject is a 70 a.D. travertine amphitheatre whose creator (as contractor) was Vespasianus. If we add the issue of the media type of the web resource, as introduced in 3.1, an additional level becomes manifest: we are describing a 21st century JPEG image of a 20th century photograph of a 18th century print of a 1st century building.

One of the most frequent dilemmas for a provider of metadata about an object is deciding whether interesting information for which no natural facet is available should be omitted, forced into an inappropriate facet (e.g., the *dc:type* or *dc:date* in the above examples), or dumped into a generic container (e.g., the *dc:description* above). A better solution would be to use a metadata model whose characteristic naturally accommodates the interesting information. As such, a simple solution exists already for the layers of subjects: while in DC the subject is "the topic of the resource [that is] typically […] represented using keywords, key phrases, or classification codes", in FRBR "the «has as subject» relationship indicates that any of the entities in the model, including work itself, may be the subject of a work".

Thus the example of the print by Piranesi could be expressed more precisely and with fewer misunderstanding as a record for a JPEG image whose *frbr:subject* is a 1956 photo whose *frbr:subject* is a 1756 print whose *frbr:subject* is a roman building, for instance as in figure 1[15]:

```
ontology:Photography rdfs:subClassOf frbr:Work .
ontology:Print rdfs:subClassOf frbr:Work .
ontology:Amphitheatre rdfs:subClassOf frbr:Work .

resource:jpeg-photo a ontology:JPEGImage
     ; frbr:subject resource:photo .

resource:photo a ontology:Photography
     ; dc:date "1956"
        ; frbr:subject resource:antichità-romane .
```

Fig. 1. An OWL rendering of the correct relationships between subject layers using FRBR (all other facets are expressed as in the original examples for simplicity)

[15] In this and all subsequent examples, we use the following prefixes (please note that the prefixes with the Europeana domain are fictitious, are present in these example only as a suggestion and do not correspond to existing ontologies):

```
@prefix resource: <http://data.europeana.eu/resource/>
@prefix ontology: <http://data.europeana.eu/ontology/>
@prefix foaf: <http://xmlns.com/foaf/0.1/>
@prefix pro: <http://purl.org/spar/pro/>
@prefix dc: <http://purl.org/dc/elements/1.1/>
@prefix frbr: <http://purl.org/vocab/frbr/core#>
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
resource:antichità-romane a ontology:Print
     ; dc:date "1756"
       ; dc:creator "Giovanbattista Piranesi"
       ; frbr:subject resource:colosseum .
resource:colosseum a ontology:Amphitheatre
     ; dc:date "70/80"
       ; dc:creator "Vespasianus [Auftrag]" .
```

Fig. 1. *(Continued)*

Using FRBR in its true meaning, so as to distinguish the stratification of layers resulted from describing an object (the idea, the content, the format and the specific item), it is also possible to better distinguish between the different levels (the work, the expression, the manifestation and the item).

EDM, in fact, defines *dc:subject* more precisely than Dublin Core itself, explicitly specifying that its value is either a string or a reference (thus allowing references to other CHOs), and even defines a subproperty of *dc:subject*, called *edm:isRepresentationOf*, to precisely specify the relationship between representations and represented entities (e.g., a statue and a painting of the statue). Yet, the *edm:isRepresentationOf* property is currently greyed out (meaning that it "will not be used in the first implementation so any values provided for them will not be used"), and the current number of subjects specified as strings will make turning them into references a conspicuous and non-trivial job[16].

3.3 Roles and Values

Many of DC properties (e.g., *dc:creator* and *dc:contributor*) are considered insufficient so often that in most Europeana resources that we have checked many actual values are composed of the indication of a role or other contextual information as well as a name (e.g. of the creator and/or contributor). For example consider:

Creator: Morel, Francois (Radierer) [17]
Creator: Cartographer : Ryther, Augustus[18]

[16] …not to mention the fundamental problem that in OWL a property can either be a data property (i.e., allowing strings) ***or*** an object property (i.e., allowing references) but not both, so that any OWL ontology based on EDM will have to choose one representation for the values of all the properties, including *dc:subject*, that allow either strings or references as their values. This, in and by itself, will be a major exercise in reconversion and qualification of existing data sets.

[17] http://www.europeana.eu/portal/record/08547/DC2A5E3DB3A0675D12DA7699647D1D8FA1B9293D.html

[18] http://www.europeana.eu/portal/record/92037/25F9104787668C4B5148BE8E5AB8DBEF5BE5FE03.html

Creator: Friedrich, J. C. F. [Production][19]
Subject: AUTN=Piranesi Giovanni Battista; AUTA=1720/ 1778[20]

The universality of this approach is evident, and similarly evident is the need to provide more information than a mere name, although the syntaxes, the metadata models and the provided information differ.

A better solution could be obtained by promoting strings into first-class objects – e.g., converting people names into individuals of the class *edm:Agent* or *foaf:Person*[21] – and dealing with people names and people's roles separately. There exist two alternative ways to address efficiently and effectively this issue.

On the one hand, we can create explicit sub-properties of properties such as *dc:creator* or *dc:contributor*. For instance, by allowing "cartographer" to become an explicit sub-property (e.g., property *ontology:cartographer*) of *dc:creator*, the identification of name and roles becomes possible and easy, and consequently the queries become more powerful, as shown in the following excerpt related to *The Cittie of London 31* by Augustus Ryther:

```
resource:cittie-of-london-31 ontology:cartographer resource:ryther .

resource:ryther a foaf:Person
     ; foaf:givenName "Augustus"
     ; foaf:familyName "Ryther" .
```

A problem with this approach is that the TBox of the ontology needs to be modified every time a new role is defined as a new subproperty of *dc:creator,* which is not a good design principle in general.

An alternative is to define people's roles as individuals of a class. In theory, CIDOC CRM already implements this behaviour by using the meta-property *P14.1 in the role of* [13] (a property of property *P14 carried out by*) so as to specify the role that an agent has in the context of a particular event (e.g., the creation of an artistic work) through an instance of the class *E55 Type*. However, the official RDFS ontology of CIDOC CRM[22] does not implement any meta-property and in reality RDF lacks the expressive power needed to define meta-properties altogether. To simulate a meta-property in RDFS/OWL we may define an additional class that associates the person to his/her role. For instance, the *Publishing Roles Ontology* (*PRO*)[23] has this behaviour by means of the class *pro:RoleInTime*,:

[19] http://www.europeana.eu/portal/record/08547/AD78BAEF3D932EF43765BAD78FE8E707EA4AFF85.html
[20] http://www.europeana.eu/portal/record/08504/3F54955AB672A4DA3C5A9C268D659A0075170C7F.html
[21] http://xmlns.com/foaf/spec/#term_Person
[22] http://www.cidoc-crm.org/rdfs/cidoc_crm_v5.0.4_english_label.rdfs
[23] http://purl.org/spar/pro

```
resource:ryther a foaf:Person
     ; foaf:givenName "Augustus"
     ; foaf:familyName "Ryther"
     ; pro:holdsRoleInTime [ a pro:RoleInTime
            ; pro:withRole resource:cartographer
              ; pro:relatesTo resource:cittie-of-london-31 ] .

resource:cartographer a pro:Role .
```

This approach has the advantage of not requiring the modification of the TBox of the ontology whenever a new role is needed: we have just to add a new individual of the class *pro:Role*.

4 Experimenting on "data.europeana.eu"

The current implementation of the web site `http://data.europeana.eu` already contains a first selection of the full library of items as RDF statements and they are already queryable via Linked Data aggregators[24]. However some limits can be observed in how the site handles non-existing and non-translated resources, which prevents this implementation from being fully compliant with the Linked Data architecture.

According to [15], one of the most important principles of Linked Data is that all the published resources must be defererenceable. *Content negotiation* is necessary, since information about a resource should be always returned according to the format requested by the user who is navigating the Linked Data, e.g., HTML for humans and Turtle for computer agents. Content negotiation usually has the form of a "303 redirect": the client asks for a resource in a specific format, the server answers with a "303 See Other" HTTP status code indicating the URL where that requested representation is available to the client, and finally the client gets the content from the specified URL.

Europeana does in fact correctly implement the "303 redirect" approach for the resources it makes available in RDF, but does not behave correctly for non-existing or non-available resources. Regardless of whether the resource exists or not, in fact, the server always returns a 303 redirect, and then, after the client restates the query to the new URL, it returns an error if the resource is non-existent. Good Linked Data policy, on the other hand, is that 303 is only returned on existing resources, and an immediate error is returned for non-existing or non-available resources.

Two different approaches can be adopted for the return code, depending on which perspective is adopted: in an *Open World perspective,* we cannot state whether a resource exists, but we can only say whether we have data about it, while in a *Closed World perspective*, if no data is available about a resource, then the resource itself does not exist.

[24] Although the Europeana Linked Data project is still ongoing, we hope that what we describe in this section can be seen as valuable and meaningful suggestions for future modifications of the Linked Data infrastructure of Europeana.

Depending on which of the above views the server adopts, the client should expect a different reply than a "303 See Other" when its initial request cannot be satisfied, as shown in table 1.

Table 1. HTTP status code for non-existing or non-available resources in Linked Data

	Open World view	**Closed World view**
Resource is not available in the requested format	*406 Not Acceptable* Information about the resource exist but they cannot be returned in the requested format	*406 Not Acceptable* There are no information about the resource in the requested format
There is no resource with that URI	*204 No Content* (no body specified) or *303 See Other* and the indication of a new location that contains empty content in the format requested, e.g. an empty RDF/XML string "<*rdf:RDF* />" There are no information about the resource	*404 Not Found* The resource does not exist

5 Conclusion

The final question is: how can we improve user queries? How do we work in the direction of an effective enrichment of metadata, in order to address the information needs of the end users? Europeana currently misleads in the object descriptions mainly because of the imprecise mapping of the original metadata set onto the Europeana specific model. The variety of metadata vocabularies, ontologies and models makes things difficult to manage. Approaches towards a better integration of the different metadata sources that feed Europeana could be helped by existing works on the creation, extension and alignment of OWL ontologies, but much work in the mapping of richer models still needs to be dealt with by hand. The EDM Mapping Guidelines [13] should lead content providers to create descriptions compliant to EDM, at the same time leaving them free to use metadata models and value vocabularies most appropriate to their own internal uses. But although by correctly using the Guidelines many of the existing problems would be solved, some aspects of the EDM could be improved, reflecting on the different levels of description of the objects. Full and correct Linked Data compliancy, furthermore, is the right direction for the future and will help Europeana in giving more complete and structured descriptions. Yet, the techniques have to be refined. We wait for the announced Europeana v2.0 at the end of 2014[25].

References

1. Aloia, N., Concordia, C., Meghini, C.: Europeana v1.0. In: Agosti, M., Esposito, F., Meghini, C., Orio, N. (eds.) IRCDL 2011. CCIS, vol. 249, pp. 127–129. Springer, Heidelberg (2011)
2. Europeana Project. Definition of the Europeana Data Model. Version 5.2.3 (February 24, 2012), `http://pro.europeana.eu/documents/900548/bb6b51df-ad11-a78-8d8a-44cc41810f22`

[25] `http://pro.europeana.eu/web/europeana-v2.0`

3. Europeana Project. Europeana Semantic Elements Specification, Version 3.4 (March 31, 2011), http://pro.europeana.eu/documents/900548/4968d0bd-416b-48ed-bc67-6a4a47f09098
4. Europeana Project. Europeana Data Model Primer (October 26, 2011), http://pro.europeana.eu/documents/900548/770bdb58-c60e-4beb-a687-874639312ba5
5. Haslhofer, B., Isaac, A.: data.europeana.eu: The Europeana Linked Open Data Pilot. In: DCMI Proceedings of the International Conference on Dublin Core and Metadata Applications, pp. 94–104 (2011)
6. Haslhofer, B., Roochi, E.M., Gay, M., Simon, R.: Augmenting Europeana Content with Linked Data Resources. In: Paschke, A., et al. (eds.) Proceedings of the I-Semantics, 6th International Conference on Semantic Systems. ACM (2010)
7. Dublin Core Metadata Initiative. Dublin Core Metadata Element Set, Version 1.1. DCMI Recommendation (2010), http://dublincore.org/documents/dces/
8. Dublin Core Metadata Initiative. DCMI Metadata Terms, DCMI Recommendation (2010), http://dublincore.org/documents/dcmi-terms/
9. Miles, A., Bechhofer, S.: SKOS Simple Knowledge Organization System Reference. W3C Recommendation (August 18, 2009), http://www.w3.org/TR/skos-reference/
10. International Federation of Library Associations and Institutions. Functional Requirements for Bibliographic Records Final Report (2009), http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf
11. Lagoze, C., Van de Sompel, H., Johnston, P., Nelson, M., Sanderson, R., Warner, S.: Abstract Data Model. ORE Specification. Open Archives Initiative (October 17, 2008), http://www.openarchives.org/ore/1.0/datamodel
12. TC 46/SC 4. ISO 21127:2006. Information and documentation–A reference ontology for the interchange of cultural heritage information. International Organization for Standardization (2006)
13. Crofts, N., Doerr, M., Gill, T., Stead, S., Stiff, M.: Definition of the CIDOC Conceptual Reference Model. Version 5.0.4, ICOM/CIDOC CRM Special Interest Group (November 2011), http://www.cidoc-crm.org/docs/cidoc_crm_version_5.0.4.pdf
14. Europeana Project. Europeana Data Model Mapping Guidelines (February 24, 2012), http://pro.europeana.eu/documents/900548/ea68f42d-32f6-4900-91e9-ef18006d652e
15. Heath, T., Bizer, C.: Linked Data: Evolving the Web into a Global Data Space. Morgan & Claypool Publishers (2011), doi:10.2200/S00334ED1V01Y201102WBE001