EAC-CPF Ontology and Linked Archival Data

Silvia Mazzini[1], Francesca Ricci[2]

[1] Regesta.exe (Rome, Italy)
smazzini@regesta.com
[2] Istituto per i beni artistici culturali e naturali della Regione Emilia-Romagna (IBC) (Italy)
fricci@regione.emilia-romagna.it

Abstract. The EAC-CPF standard is an XML schema maintained by the Society of American Archivists in partnership with the Berlin State Library used for encoding contextual information about persons, corporate bodies, and families related to archival materials. The main goal of this paper is to demonstrate the feasibility of the application of Semantic Web technology for creating Linked Open Data of descriptions of entities associated with the creation and maintenance of archives. In this paper we present two EAC-CPF ontologies and we provide an in-depth description of all phases of the work, from the study of the standard to the definition of the classes and properties of the two OWL ontologies and a case study of application in authority records of *IBC Archivi* (information system of historical archives in the Emilia-Romagna region).

Keywords. EAC-CPF, ontology, RDF, Linked Open Data, archival description.

1 Introduction

International standards for archival and encoding descriptions are known for a long time in Italy. EAD (Encoded Archival Description)[1] standard has been introduced in Italy early in public and private area, so today many archival description software use EAD schema or offer an XML export for the resources. By now, XML[1] is known as a good standards for semantic interoperability and it is often used for representing archival resources thanks to its simplicity, its flexibility and its capabilities of nesting description particularly useful to archival multi-level description.

Furthermore semantic interoperability is a *sine qua non* for the Semantic Web and today archivists have to deal with the nascent Semantic Web. It is now quite common to use links as means of connecting archival descriptions on the web to other information, in order to increase the information available to users who access archival material on the web.

Increasing development of Linked Open Data in cultural heritage leads to a review of technologies in other areas too, like e.g. the archival domain. We believe that technologies that best introduce archival description background to web of data are RDF [3] and ontologies [4]. In addition to these reasons, we can say that behind the

[1] http://www.loc.gov/ead/

idea to transform the EAC-CPF schema into an ontology and the experiment to "open" eac-cpf authority records as linked open data, there are also:

- the need to describe the resources in a format that can be shared and approved by the international scientific community;
- the choice to use standards allows to process, integrate and deal with data according to standardized rules that are supported by large communities;
- the opportunity to integrate with other web resources described with other standard vocabularies.

Starting from these considerations, we believe that a concrete solution is to use RDF and ontologies, not only as means for representing entities and the relations between the various components of the archival description, but also as an appropriate tool to qualify these relations semantically.

A few simple actions are required to be done in order to describe archival context in a "semantic way". It is necessary to:

1. identify univocally the descriptive resources by means of the URI and preferably use dereferenceable URI;
2. provide descriptions in a standard format so that the resources and their relations can be recognized immediately;
3. include in the descriptions the greatest possible number of relevant links to other information resources.

The current digital environment is clearly oriented towards a more intelligent web, able to support the sharing, enhancement and management of archival information, exploring the meaning of the documents and returning data (and not documents).

Linked Data[2] and ontologies are the technological components on which the passage from Web 2.0 to the Semantic Web is based. However, to make this change a reality, the technological components are not sufficient but it will be necessary for those who publish data on the web to do so in a "open" way, thus contributing to the realization of a truly "open" semantic web.

On the basis of these first premises, the Istituto per i beni artistici culturali e naturali (IBC) of the Emilia-Romagna Region has decided to open up its archival data.

IBC was founded in 1974 and it's the scientific and technical instrument for the Emilia-Romagna regional planning in the field of artistic, cultural and environmental heritage. The Soprintendenza regionale per i beni librari e documentari has been part of IBC since 1983, with the specific task of co-ordinating the regional policy addressed to libraries and archives[3] .

IBC develops the IT facilities that convey archives, libraries and museums data to institutions and the general public, promotes and coordinates the census and the description of archival, book and museum material, grants the readability of specific

[2] http://linkeddata.org/

[3] http://www.ibc.regione.emilia-romagna.it/wcm/ibc/pagine/01chi_inglese.htm

DBs on the web and at present IBC's working on the standards for interoperability through the use of semantic web technologies.

In March 2001 a group of archivists met in Toronto and created a high-level model for the description of individuals, families and corporate bodies that create, preserve, use and are responsible for and/or associated with archival records in a variety of ways. The group has termed the model "Encoded Archival Context - Corporate Bodies, Persons, and Families" (EAC-CPF)[4] to emphasize its important role in archival description and its relationship with the Encoded Archival Description standard.

Since the EACWG meeting in Bologna and the conference "Standards and exchange formats for interoperability among archival information systems" organized by IBC in early May 2008[5], IBC has been committed to the dissemination of EAC-CPF in the Italian context, to promoting knowledge and use of this standard by Italian archivists and archival agencies and to translate in Italian the EAC-CPF tag library[6].

The first step in this direction was the opening of a standard (by publishing an ontology for EAC-CPF in an open format and including parts of other standards within it). Afterwards a second ontology was realized to represent the EAC-CPF records containing the descriptions of archival creators published in *IBC Archivi* (information system of historical archives in the Emilia-Romagna region)[7]. These two ontologies are complementary and closely related because the experience with devising the first one has provided the basis to define the approach for devising and using the second one. In this paper we present:

- the first ontology (described in chapter 2) that is a different formalization of the XML schema of EAC-CPF standard, useful to promote and foster a better comprehension of structure and properties of the standard among Italian archivists;
- the second ontology (described in chapter 3) that was realized to open -by the semantic web- the descriptions of entities (corporate bodies, persons and families) associated with the creation and maintenance of archives;
- an example realized on *IBC Archivi* descriptions (described in chapter 4).

2 EAC-CPF standard Ontology

The EAC-CPF Schema has a fairly simple structure with much less nesting than its relative for archival description EAD: specifies 90 elements and 30 attributes[8]. The structure is designed in such a way as to maintain a division between information controlling the entity and its analytic description.

---

4 http://eac.staatsbibliothek-berlin.de/about.html

5 http://online.ibc.regione.emilia-romagna.it/h3/h3.exe/apubblicazioni/sD:!TEMP!HwTemp!3se2a84aa31d.tmp/d1/FFormDocumento?La.x=;sel.x=NRECORD%3d0000047818

6 IBC entrusted the italian translation of EAC-CPF tag library to Salvatore Vassallo, under the scientific supervision of Stefano Vitali.

7 http://archivi.ibc.regione.emilia-romagna.it/ibc-cms/

8 http://eac.staatsbibliothek-berlin.de/eac-cpf-schema.html

Following an analysis of the relations between elements of the schema and attributes, we thought of proceeding to a first semantic web description of the schema (using OWL) by aiming to create a different formalization of the EAC-CPF standard, to provide a new tool for navigating the schema showing the relations, and pointing to specifications of the official tag library and the diagram of the xml schema for the technical specifications of each element.

The XML schema of EAC-CPF does not present much nesting in the description and, it was fairly simple to convert it into OWL ontology without changing the general settings of the standard and without introducing any new elements. In general, the RDF data model is based on the official schema of EAC-CPF standard. It is not proposed as an alternative standard but quite simply as a different formulation, which is useful for the semantic web and fosters interoperability.

2.1 Classes and properties of the EAC-CPF standard ontology

The first ontology describes strictly the domain of the XML schema so that we have created only three *owl classes* (element, attribute and controlled_value) and few properties useful to represent schema's relations.

Table 1. class and properties of ontology

Classes: element, attribute, controlled_value. **Properties**: mayContainElement, containRequiredElement, hasAttribute, hasRequiredAttribute, mayContainValue, reference, isElementOf, isRequiredElementOf, isAttributeOf, isRequiredAttributeOf, isControlledValueOf, mayContainDatatype, diagram_ref, occurrence.

Fig. 1 shows an RDF serialization of description of identity element based on the ontology. URIs for the resources are URLs of the element in the tag library official web site.

```
<eac-cpf:element rdf:about="http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html#d1e3528">
    <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string" xml:lang="en">&lt;identity&gt;</rdfs:label>
    <dc:title xml:lang="en">Identity</dc:title>
    <dcterms:abstract xml:lang="en">A wrapper element for the name portion of the EAC-CPF instance.</dcterms:abstract>
    <dc:description xml:lang="en">&lt;identity&gt; is a container element grouping all the necessary elements to the name
identification of the EAC-CPF instance. Within this element the &lt;entityType&gt; element is required and specifies the type of
entity being described (i.e., corporateBody, family, or person), and one or more &lt;nameEntry&gt; elements and / or one or
more &lt;nameEntryParallel&gt; elements specifying names by which the entity is known. An optional &lt;entityId&gt; is available
for any identifiers associated with the entity. All the names by which the entity being described in the EAC-CPF instance is
known are provided within this element. Within &lt;identity&gt;, the names of the entity, whether authorized or alternatives,
should be recorded in a &lt;nameEntry&gt; element. If there is more than one name for the entity, each of them should be
recorded in a separate &lt;nameEntry&gt; element. In addition to needing to accommodate one or more names used for or by the
entity, &lt;identity&gt; may accommodate two or more parallel names in different languages or scripts. In countries where there
is more than one official language, such as Canada or Switzerland, names of entities are frequently provided in more than one
language. Within &lt;identity&gt;, a &lt;nameEntryParallel&gt; element should be used to group two or more &lt;nameEntry&gt;
elements that represent parallel forms of the name of the entity being described. Within &lt;identity&gt;, a &lt;descriptiveNote&gt;
element may be used to record other information in a textual form that assists in the identification of the entity.    </
dc:description>
    <eac-cpf:mayContainElement rdf:resource="http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html#d1e2600"/>
    <eac-cpf:mayContainElement rdf:resource="http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html#d1e2711"/>
    <eac-cpf:containRequiredElement rdf:resource="http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html#d1e2773"
/>
    <eac-cpf:mayContainElement rdf:resource="http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html#d1e5245"/>

    <eac-cpf:hasAttribute rdf:resource="http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html#identityType"/>
    <eac-cpf:hasAttribute rdf:resource="http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html#localType"/>
    <eac-cpf:hasAttribute rdf:resource="http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html#xmlbase"/>
    <eac-cpf:hasAttribute rdf:resource="http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html#xmlid"/>
    <eac-cpf:hasAttribute rdf:resource="http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html#xmllang"/>
    <eac-cpf:reference rdf:resource="isaar.rdf#5.1"/>
    <eac-cpf:diagram_ref>http://eac.staatsbibliothek-berlin.de/Diagram/cpf.html#id83</eac-cpf:diagram_ref>
    <eac-cpf:occurrence>1</eac-cpf:occurrence>
</eac-cpf:element>
```

Fig. 1. RDF serialization of Identity element

The graph below (fig. 2) shows a visualization of the same element (*identity*) of the standard, its relations with other element of the schema (orange circles) and with attributes (yellow circles); while the color of arrows and the direction clarify the type of relation.

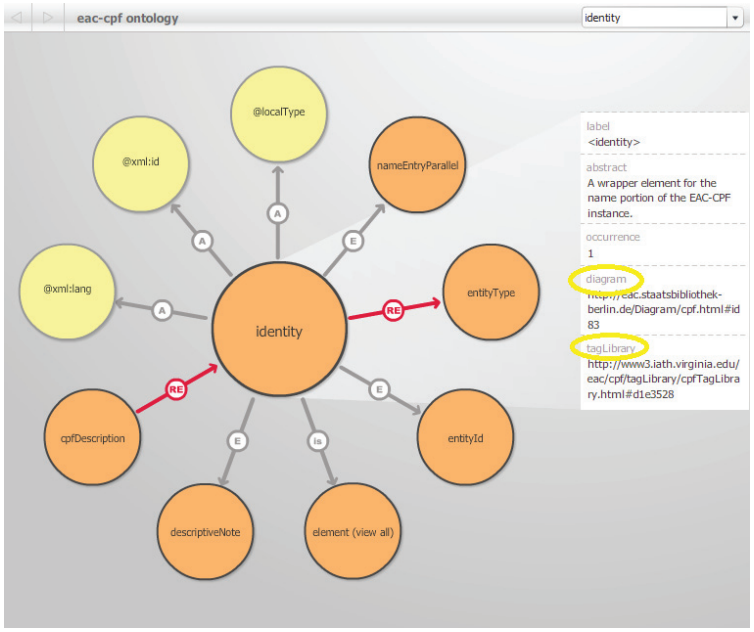

Fig. 2. Graph visualization of identity element in Relation Browser

This initial study was concluded last summer with publication of ontology and graph visualization on the web site of the Libray Linked Data Incubator Group[9].

3 EAC-CPF Descriptions Ontology for Linked Archival Data

The work described in chapter 2 was extremely useful as a feasibility study and an effective work tool for archivists, but it could not be used to open the authority records codified with this standard to the world of Linked Open Data. It was necessary to transform the elements of the schema into properties of the ontology and to change the point of view of the description of the model. It was necessary to move from the description of the XSD schema in RDF to the definition of a new model based on the schema (thus maintaining the names of the elements and the attributes). For example, if you write a text in the EAC-CPF tag <bioghist> of an XML file, you mean that the text is a "history of the institution" or a "biography". If you want to obtain the same result in an RDF file, you have to change the xml element <bioghist>

[9] http://www.w3.org/2005/Incubator/lld/wiki/Vocabulary_and_Dataset

into the RDF property eac-cpf:bioghist. In this way, you assign a semantic value to the text itself.

To reach a description of the data model (that could be used for the Linked Archival Data), it was necessary to take a further step: starting from the records describing the authorities, bodies, persons and families of the *IBC Archivi* codified in EAC-CPF, we moved on to the definition of a data model based on the standard, maintaining the names of the elements and the attributes and the relations, but expressing them in RDF. In general, the following basic principles were followed:

- to make the RDF model more explicit, the three typologies of entities (that are included in EAC-CPF schemas as control values for <entityType> element), have become three distinct classes in the ontology: Person, Family, and Corporate Body as subclasses of the more general Entity;
- no new concepts have been added that were not defined in the XML schema;
- if the standard proposes the names of the elements in both the singular and the plural form, in the RDF data model only the singular forms have been maintained, since properties can always be repeated in RDF;
- the elements used in the XML schema to parcel the descriptive information were not used in the data model, aiming to group the information favouring a simpler and more general structure. For example, the element <p> present in almost all the descriptive elements was omitted, as well as the formatting elements (such as span, list, item, level, outline, etc.);
- in the RDF file some information, especially classical descriptive metadata such as title, date and author were duplicated by using other RDF terms that are universally known and used such as Dublin Core and FOAF to allow a natural interoperability with other similar resources;
- to facilitate the linking of external resources and build up the linked archival data, for all those resources for which it was possible to find alternative URIs or alternative information on other websites or with other authorities, the references were added: for example, to link the names of persons to the Virtual International Authority File (VIAF), we used a property of OWL owl:sameAs since this indicates that two URI references actually refer to the same thing - the individuals have the same "identity". The same is true for names of places of birth and death, the property eac:place is not an xml Literal but the URI of a place described in GeoNames database.

3.1 Classes and properties of the ontology

The EAC-CPF schema is made up of two macro sections in which the record control information and the metadata descriptors converge. Therefore in order to reproduce this situation in the EAC-CPF ontology, we created the class *controlArea* and the class *descriptionArea* which contain all the specific information.

The relations between other entities or other resources are managed by a class *relation* which directly points either to other URIs or to resources outside the system.

We introduced the following classes and properties:

Table 2. class and properties of ontology

Classes: entity, person, corporateBody, family, controlArea, descriptionArea, nameArea, language, place, relation. **Properties**: authorizedForm, biogHist, control, conventionDeclaration, cpfRelation, cpfRelationType, description, existDates, function, generalContext, languageDeclaration, languageUsed, legalStatus, localTypeDeclaration, maintenanceAgency, maintenanceHistory, maintenanceStatus, mandate, nameEntry, occupation, publicationStatus, recordID, resourceRelation, resourceRelationType, source, structureOrGenealogy

Basically, the graph obtained by the proposed ontology is the following:

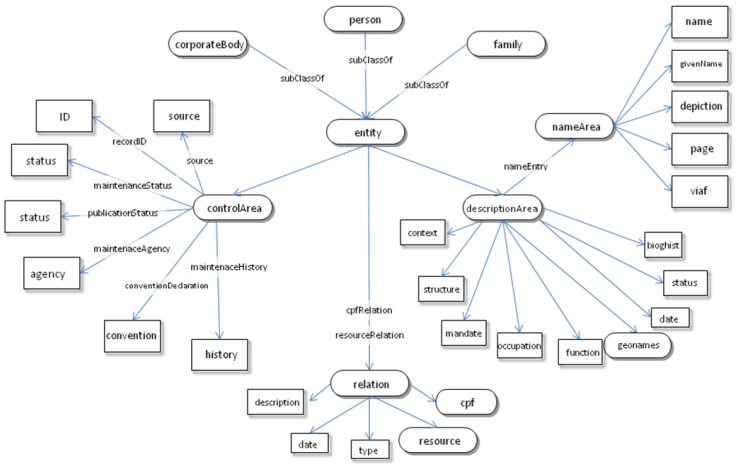

Fig. 3. Graph of the ontology

3.2 External RDF vocabularies references

As far as possible, we have tried to make use of the other popular and widely accepted and supported RDF vocabularies that already exist in the field of cultural heritage and generally in the world of linked data. Besides the Semantic Web languages OWL, RDF and RDFS, we also used the vocabularies: *skos*[10] – Simple Knowledge Organization System, *foaf*[11] – Friend of a Friend, *dc*[12] – Dublin Core, *Bio*[13] - biographical ontology, *Viaf*[14] - The Virtual International Authority File, *Gn*[15] – GeoNames.

[10] http://www.w3.org/2004/02/skos/core#
[11] http://xmlns.com/foaf/0.1/
[12] http://purl.org/dc/elements/1.1/ and http://purl.org/dc/terms
[13] http://purl.org/vocab/bio/0.1/
[14] http://viaf.org/ontology/1.1/#
[15] http://www.geonames.org/ontology#

4 Example

For many years IBC has been experimenting with archival description standards and encoding systems for describing archival institutions, historical archives and creators in the Emilia-Romagna region; actually in *IBC Archivi* the descriptions of 389 archival institutions, 2230 historical archives and 185 creators are published.

This is why we tried to imagine a network (or a graph) which expands slowly but progressively. The graph could show all the resources dynamically connected to it: both the *IBC Archivi* descriptions and the descriptive data opened by other systems and similar environments (libraries, museums, cultural institutions in general, etc.) and recovered thanks to the semantic network.

For example we imagined a map of the Emilia-Romagna region which shows the location of the archival institutions described in the *IBC Archivi*. If we use the GeoNames ontology to reference the institutions locations, automatically the institutions and their archives will be connected to all the other resources referenced in the same place through GeoNames.

In this first test phase, the field of application chosen for this project is the set of descriptive files of the archive producers created in the context of the *IBC Archivi* information system. The authority records of archive producers (about 400, including corporate bodies, persons and families, described in EAC-CPF format) were created on the IBC-xDams platform (a web-based platform for EAD and EAC compliant archive file creation). This is why these descriptions constitute the project's testbed.

A first example was made with the authority record "Andrea Costa"[16], whose papers are kept at the municipal historical archives of Imola and are described using the IBC-xDams platform[17]. The "Andrea Costa" record, in particular, is a suitable case study because it has a fairly analytic description and numerous relations with other archive producers described and with various typologies of resource contained in *IBC Archivi* and in other information systems.

We tried to read the RDF files produced in this way (fig. 4) with an open source faceted browser called *Longwell*[18] created for the Simile project[19]. Faceted navigation adapts well to RDF files precisely because they are not hierarchical files but there only transverse relations between the resources and so it is easy to visualize the data from different points of view or facets; at the same time it is possible to set and remove filters, derived from the properties introduced into the ontology, which allow navigation to be guided and targeted. In this Longwell faceted browser there are some additional small features thanks to the resources which are connected in the RDF. It is possible to visualize on the map the locations that the browser recognizes as such simply because they have already been identified with *GeoNames'* URI and to obtain a graph that best expresses the relations between the resources.

[16] Andrea Costa (Imola 1851-1910) was an Italian socialist activist, he was born in Imola and he co-founded the Partito dei Lavoratori Italiani in 1892

[17] http://www.regesta.com/cosa-e-xdams/

[18] http://simile.mit.edu/wiki/Longwell

[19] http://simile.mit.edu/

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"...>
        <eac-cpf:person rdf:about="http://archivi.ibc.regione.emilia-romagna.it/eac-cpf/IT-ER-IBC-SP00001-0000264">
                <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Andrea Costa, 1851 - 1910</rdfs:label>
                <foaf:page rdf:resource="http://dbpedia.org/page/Andrea_Costa"/>
                <dc:date>18511129-19100119</dc:date>
                <foaf:depiction rdf:resource="http://fondazione.camera.it/sites/default/files/1-andrea_costa_a_100_anni_dalla_scomparsa_big_b_0.jpg"/>
                <owl:sameAs rdf:resource="http://viaf.org/viaf/62357877/"/>
                <gn:locationMap rdf:resource="http://www.geonames.org/3175537/"/>
                <eac-cpf:control rdf:parseType="Resource">
                        <dc:title>Identificazione Soggetto Produttore Andrea Costa</dc:title>
                        <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Identificazione Andrea Costa</rdfs:label>
                        <eac-cpf:recordID rdf:resource="IT-ER-IBC-SP00001-0000264"/>
                        <eac-cpf:maintenanceStatus>new</eac-cpf:maintenanceStatus>
                        <eac-cpf:publicationStatus>approved</eac-cpf:publicationStatus>
                        <eac-cpf:maintenanceAgency>Scheda descrittiva a cura di ...</eac-cpf:maintenanceAgency>
                        <eac-cpf:languageDeclaration rdf:resource="http://id.loc.gov/vocabulary/iso639-1/it"/>
                        <eac-cpf:conventionDeclaration>IBC</eac-cpf:conventionDeclaration>
                        <eac-cpf:source>Il profilo biografico di Andrea Costa è desunto dalla voce Andrea Costa di ...</eac-cpf:source>
                </eac-cpf:control>
                <eac-cpf:description rdf:parseType="Resource">
                        <eac-cpf:nameEntry rdf:parseType="Resource">
                                <foaf:name>Andrea</foaf:name>
                                <foaf:givenName>Costa</foaf:givenName>
                                <foaf:depiction rdf:resource="http://fondazione.camera.it/sites/default/files/1-andrea_costa_a_100_anni_dalla_scomparsa_big_b_0.jpg"/>
                                <owl:sameAs rdf:resource="http://viaf.org/viaf/62357877/"/>
                                <foaf:page rdf:resource="http://en.wikipedia.org/wiki/Andrea_Costa"/>
                                <foaf:page rdf:resource="http://dbpedia.org/page/Andrea_Costa"/>
                        </eac-cpf:nameEntry>
                        <eac-cpf:existDates>18511129-19100119</eac-cpf:existDates>
                        <bio:Birth rdf:parseType="Resource">
                                <bio:date>29 novembre 1851</bio:date>
                                <bio:place rdf:resource="http://sws.geonames.org/3175537/"/>
                        </bio:Birth>
                        <bio:Death rdf:parseType="Resource">
                                <bio:date>19 gennaio 1910</bio:date>
                                <bio:place rdf:resource="http://sws.geonames.org/3175537/"/>
                        </bio:Death>
                        <eac-cpf:biogHist>biografia di Andrea Costa...</eac-cpf:biogHist>
                        <eac-cpf:cpfRelation rdf:parseType="Resource">
                                <eac-cpf:cpfRelationType>associative</eac-cpf:cpfRelationType>
                                <dc:description>Andrea Costa è legato a Serafino Mazzotti ...</dc:description>
                                <dcterms:relation rdf:resource="http://archivi.ibc.regione.emilia-romagna.it/eac-cpf/IT-ER-IBC-SP00001-0001003"/>
                                <dc:date>1874 - 1910</dc:date>
                        </eac-cpf:cpfRelation>
                        <eac-cpf:cpfRelation rdf:parseType="Resource">
                                <eac-cpf:cpfRelationType>associative</eac-cpf:cpfRelationType>
                                <dc:description>Andrea Costa è legato ad Anna Kuliscioff da un intenso ...</dc:description>
                                <dcterms:relation rdf:resource="http://archivi.ibc.regione.emilia-romagna.it/eac-cpf/IT-ER-IBC-SP00001-0001004"/>
                                <dc:date>1877 - 1910</dc:date>
                        </eac-cpf:cpfRelation>
                        <eac-cpf:resourceRelation rdf:parseType="Resource">
                                <eac-cpf:resourceRelationType>creatorOf</eac-cpf:resourceRelationType>
                                <dc:description>soggetto produttore</dc:description>
                                <dcterms:relation rdf:resource="http://archivi.ibc.regione.emilia-romagna.it/ead-str/IT-ER-IBC-AS00209-0004118"/>
                                <dc:date>1872 - 1910</dc:date>
                        </eac-cpf:resourceRelation>
                        <eac-cpf:resourceRelation rdf:parseType="Resource">
                                <eac-cpf:resourceRelationType>other</eac-cpf:resourceRelationType>
                                <dc:description>possessore</dc:description>
                                <dcterms:relation rdf:resource="http://archivi.ibc.regione.emilia-romagna.it/ead-str/catalogo_on_line"/>
                                <dc:date>1851 - 1910</dc:date>
                        </eac-cpf:resourceRelation>
                        <skos:changeNote rdf:parseType="Resource">
                                <dc:date>12/10/2010</dc:date>
                                <dc:creator>IBC</dc:creator>
                                <rdf:value>create</rdf:value>
                        </skos:changeNote>
                        <dc:identifier>coincide con il recordID</dc:identifier>
                        <dc:creator>[AUTORE SCHEDA]</dc:creator>
                        <dc:title>Biografia di Andrea Costa e collegamenti</dc:title>
                        <dcterms:issued>2010-03</dcterms:issued>
                </eac-cpf:description>
        </eac-cpf:person>
</rdf:RDF>
```

Fig. 4. example of authority record encoded with eac-cpf ontology

5 Conclusion

The experience made with the two ontologies and the testbed on Andrea Costa's records shows that authority records can indeed be the first data to "unlock". In fact authority records by their nature are connection points between different resources. Unlocking authority record of Andrea Costa means connecting not only with his papers, but also with his library, his publications and with other related persons or entities.

We are aware that hard work still needs to be done but according to these first results, the scenario is surprising and, in particular we have to explore all the research directions. In this perspective a future collaborative effort with SNAC project[20] might be useful to share skill, tools and outcomes. At the moment we are working to build a semantic environment[21] for *IBC Archivi* in which users could utilize a SPARQL Endpoint jointly with a reasoning engine and a *linked data api* (ELDA)[22] for navigating resources.

References

1. *Extensible Markup Language (XML) 1.0 (Fifth Edition)* **W3C Recommendation** 26 November 2008, Tim Bray, Jean Paoli, C. M. Sperberg-McQueen, Eve Maler, François Yergeau eds.
2. Tim Berners-Lee: *L'architettura del nuovo Web*, Feltrinelli, 2001
3. RDF Primer, W3C Recommendation, February 10, 2004, Frank Manola , Eric Miller, eds., http://www.w3.org/TR/rdf-primer/
4. Stefan Decker et al.: The Semantic Web - on the respective Roles of XML and RDF, IEEE Internet Computing, vol.4, 2000

[20] http://socialarchive.iath.virginia.edu/index.html
[21] http://archivi.ibc.regione.emilia-romagna.it/ontology/EAC-CPF/sematic-enviroment.html
[22] http://code.google.com/p/elda/