Semantic Web 0 (2015) 1
IOS Press

Migration of a library catalogue into RDA linked open data

Gustavo Candela [a,*], Pilar Escobar [b], Rafael C. Carrasco [b] and Manuel Marco-Such [b]
[a] *Biblioteca Virtual Miguel de Cervantes, Alicante, Spain*
[b] *Departamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante, Alicante, Spain*

Abstract. The catalogue of the Biblioteca Virtual Miguel de Cervantes contains about 200,000 records which were originally created in compliance with the MARC21 standard. The entries in the catalogue have been recently migrated to a new relational database whose data model adheres to the conceptual models promoted by the International Federation of Library Associations and Institutions (IFLA), in particular, to the FRBR and FRAD specifications. The database content has been later mapped, by means of an automated procedure, to RDF triples which employ basically the RDA vocabulary (Resource Description and Access) to describe the entities, as well as their properties and relationships. Compared to a direct transformation, the intermediate relational model —ensuring, for example, referential integrity— provides tighter control over the process and, therefore, enhanced validation of the output. This RDF-based semantic description of the catalogue is now accessible online through an interface which supports browsing and searching the information. Due to their open nature, these public data can be easily linked and used for new applications created by external developers and institutions. The methods applied for the automation of the conversion, which build upon open-source software components, are described here.

Keywords:, Linked Open Data, Bibliographic and Authority Data, Cultural Heritage, Semantic Web

1. Introduction

Linked open data are defined as data which are both legally and technically interoperable: they are *open* because others are allowed to use, modify and redistribute it, and *linkable* because they are enriched with information about their relations with other data. As illustrated in figure 1, the connections between linked data are expressed as binary relations between *Uniform Resource Identifiers* (a URI is a unique identifier, such as the ISBN, used to represent objects or resources over the network), accompanied with a description of their meaning. The metadata which describe such links are often expressed in RDF (Resource Description Framework) format, a standard for data interchange[1] based on XML —see figure 2.

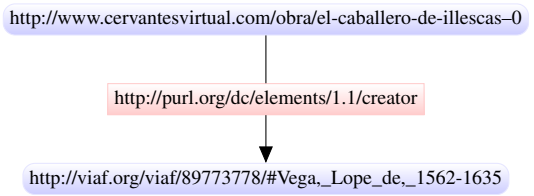

Fig. 1. An example of linked data: an online book (*El caballero de Illescas*) is connected to a VIAF personal name (Lope de Vega) through the *creator* statement, as defined by the Dublin Core standard.

Linked open data were conceived as a way to add meaning to the digital content and they are expected to support a more effective exploration and discovery of information thanks to the development of advanced search engines with deductive capacities. The benefits of publishing library information as *linked open data* include [2], among others, the following:

*Corresponding author. E-mail: gustavo.candela@cervantesvirtual.com
[1] RDF is described at http://www.w3.org/RDF.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description
     rdf:about="http://www.cervantesvirtual.com/obra/el-caballero-de-illescas--0/">
     <dc:creator rdf:resource="http://viaf.org/viaf/89773778/#Vega,_Lope_de,_1562-1635"/>
  </rdf:Description>
</rdf:RDF>
```

Fig. 2. RDF representation of the relationship between URIs depicted in figure 1.

1. Linked metadata provide a more sophisticated navigation through information and facilitates the citation and replication of experiments.
2. The enhanced interoperability increases the visibility of cultural metadata and promotes their external curation and enrichment.
3. The metadata are more durable and robust since they do not depend on a particular data structure.
4. Developers and vendors avoid being tied to library-specific formats —such as MARC.[2]

With such expectations, applying the *linked open data* concepts to the cultural heritage domain has become an active and challenging field [12]: many libraries, museums, and archives are currently exploring ways to convert their data into the RDF format, to integrate external data-sets with their own descriptions, and to develop new interfaces providing a richer experience to the users of cultural heritage websites.

In parallel, modern standards for cataloguing are emerging as an alternative replacement to the traditional ones (such as ACCR2 [1]). For example, RDA (Resource, Description and Access) is a modern vocabulary [7] for descriptive metadata supporting resource discovery. RDA follows the concepts and terminology of the Functional Requirements for Bibliographic Records (FRBR, [9]) and the Functional Requirements for Authority Data (FRAD, [14]) —and it is working to adopt the Functional Requirements for Subject Authority Data (FRSAD, [11])—, a family of models promoted by the IFLA which define entities, relationships, and attributes that should be used to describe resources.

RDA descriptions provide easier navigation and retrieval of bibliographic objects and their publication as *linked open data* create enhanced interoperability. However, the publication of bibliographic records as open data requires data preprocessing —since data are expressed primarily in natural language text and encoded using heterogeneous library standards—, as well as critical choices as regards the metadata vocabulary used to describe the library objects, the ontologies employed to specify the connections between them and the technology applied to convert large catalogues and metadata repositories.

This paper describes the steps applied for the automation and control of the migration process from a MARC21 collection of records to a set of RDF triples containing bibliographic metadata in RDA. The process relies on the creation of a relational database according to FR* conceptual models, and provides controlled generation of linked data in RDA. The implementation is strongly based on the currently available open-source technology.

2. Related work

The publication of library records as *linked open data* has been already addressed by a number of institutions. For example, the Library of Congress Linked Data Service (`id.loc.gov`) provides access to authority data such as the LC subject headings and the MARC geographic areas.

The Bibliothèque nationale de France published `data.bnf.fr` in 2011 by aggregating information about authors, works, and subjects which was scattered among various catalogues. These data are published in RDF using a vocabulary based on the FRBR model where objects are referenced through ARK identifiers.[3] The information is stored in a database which contains the data in different formats, including RDF, JSON, and HTML. [15]

The British National Bibliography Linked Data Platform (`bnb.data.bl.uk/docs`) provides ac-

[2]MARC standards are a set of digital formats for the description of items catalogued by libraries, such as books. There are several versions of MARC, the most predominant being MARC21.

[3]The *Archival Resource Key* identifiers are persistent references to web-accessible objects.

cess to the British National Bibliography (BNB), implements the SPARQL query language [19] and delivers RDF and JSON outputs. The data-set has been modelled using existing RDF vocabularies, such as Dublin Core, the Bibliographic Ontology (BIBO), and Friend of a Friend (FOAF). Exceptionally —for example, due to insufficient granularity of those vocabularies— a new term was coined and documented. FRBR was not initially used [5], since the identification of the entities in the source MARC records required extensive work. The records were therefore normalized for improved matching and later transformed into RDF using XSLT and Jena Eyeball.

The German National Library supplies its data in the RDF standard via its Linked Data Service (LDS; `http://www.dnb.de/EN/lds`) since 2010. The vocabulary is based on Dublin Core and BIBO and complemented with some elements from other vocabularies, for example, RDA, ISBD (International Standard Bibliographic Description), and GND (Gemeinsame Normdatei). The records can be also retrieved in BIBFRAME format, an RDF-based alternative to MARC21. The National Library of Spain (BNE) has recently migrated its databases to RDF and published [13] them at `datos.bne.es`. The transformation is assisted by specific software [18] —which supports RDF generation from MARC21—, and the vocabulary is strongly based on FRBR and ISBD.

The Europeana linked data at `data.europeana.eu` ensure a high level of consistency and interoperability by abstracting the original data to a common format (the Europeana Data Model). Unfortunately the richness of the original descriptions is partially lost in the homogenization process.

3. The transformation process

Traditionally, the descriptive metadata of bibliographic content —stored, for example, in MARC records— were created and interpreted by humans. Even if those records followed cataloguing rules such as AACR2 and ISBD [16], the textual descriptions therein could not be easily read and interpreted by computers —see, for instance, the rich description under field 534 in figure 3—, a common requirement in the contemporary web-connected environments. The FRBR family of conceptual models and the RDA specification provide a modern framework which facilitates the automatic processing of the information. However, the transformation of the old records into the new format has a significant cost, since libraries usually host large catalogues which should be manually revised. Therefore, software tools for the automation of the migration process are called for, and the experience of the Biblioteca Virtual Miguel de Cervantes in their implementation is described below.

```
001 ff97f774-82b1-11df-acc7-002185ce6064
003 BVC
041 $aspa
080 821.134.2-2"16"
100 $aVega, Lope de, $d1562-1635
245 $aEl caballero de Illescas
    $h[Libro electrónico]
    $c/Lope de Vega
260 $aAlicante
    $bBiblioteca Virtual Miguel de Cervantes
    $c2002
534 $aPublicación original:
    Madrid, por Juan de la Cuesta,
    a costa de  Miguel de Syles, 1620.
650 $aTeatro español $ySiglo 17º)
700 $aCuesta, Juan de la, $d 1604-1627
    $eimpresor
700 $aSiles, Miguel de $eeditor
```

Fig. 3. A MARC21 record for the novel *El caballero de Illescas*.

3.1. An FRBR-FRAD relational model for MARC21 records

A MARC21 record describes one entry in the bibliographic catalogue or authority file[4], and consists of text fields which are identified by a three-digit number —see figure 3. The text in one field can be split into sub-fields which are distinguished with a dollar sign followed by a single-character identifier. Since some fields are required (for example, field 245 containing the title) while some others are optional or user-defined, the homogeneity of the data across libraries cannot be guaranteed. Furthermore, the content of a field can be expressed with different conventions, in different languages, or it may contain typos: these features represent a challenge when MARC21 records must be shared between libraries.

The FRBR family of conceptual models [9] are intended to be independent of any cataloguing code or implementation and they identify the principal entities, their attributes and the relationships between them. The FRBR model defines the products of intellectual or artistic endeavour (work, expression, manifestation, and item) and is complemented with the FRAD model,

[4] An authority files compiles the unique terms and possible variations used to describe names, titles, and subjects.

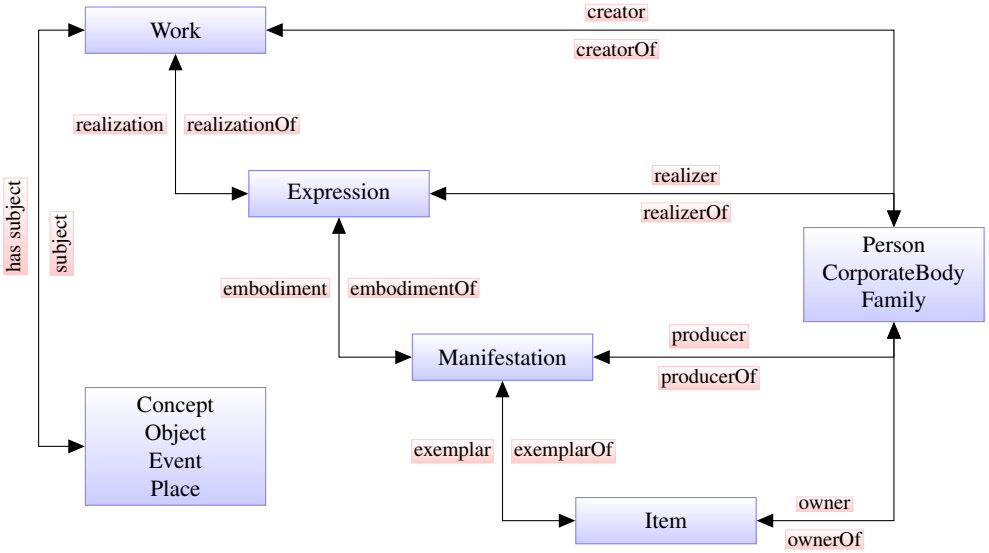

Fig. 4. Entities defined in FRBR (Work, Expression, Manifestation, Item), FRAD (Person, CorporateBody, Family), and FRSAD (Concept, Object, Event, Place) with their primary relationships.

which defines the entities responsible for the content (person, family, and corporate body), and with the FRSAD model, which defines the entities that serve as the subjects of creations (concept, object, event, and place) —see figure 4.

Inspired by the IFLA conceptual models, an Entity-Relationship (ER) model —schematically represented in figure 5— has been defined to store the Biblioteca Virtual Miguel de Cervantes descriptive metadata. Some additional elements were incorporated to the model in order to address the catalogue specificities. For example, *Collection* entities were needed to host arbitrary groupings of objects, such as works in a bibliography, items with a common provenance (e.g., a partner library holdings or items in a personal archive), which are not properly creations and usually have no associated descriptive metadata. Since authors are often the subject of a book in a library with a focus on literature, a new type of relationship was introduced to describe creations having a particular agent as *subject*; conversely, agents play different *roles* when contributing to a document —for example, printer, editor or illustrator. A generic relationship between entities (*partOf*) was defined in order to describe nested inclusions, for example, journals publishing volumes, made of issues containing articles. The online manifestations are connected to their URL with the *homepage* attribute. Entities for the UDC and Unesco classifiers and for VIAF persons were also added to the model. Since the RDA technical guidelines were created while several aspects of FRBR were still in flux, they include some additional entities (such as *Agent*) and rename some relations: for example, the FRBR *embodiment* becomes *manifestationOfExpression* in RDA –see figure 7.

As can be seen in figure 4, the abstract class *creation* generalizes the basic FRBR entities (*work*, *expression* and *manifestation*). This class has been added in order to avoid redundant descriptions (and duplicate coding), since many properties —such as subject—, are common to all types of entities.

In contrast to RDF, which does not support the validation of data structures [3], the relational model provides control over data integrity: in other approaches —for example, the XPath-based transformation described in [17]— this can be only checked programatically. For example, the referential integrity of the database will validate a *creatorOf* relationship only if the involved entities are a *work* and an *agent*; the cardinality constraints will reject a *title* which has no associated *creation*; or a *language* attribute can be defined which is only applicable to creations of type *expression*.

Once the relational database was set up, a number of tools were implemented: for example, a new catalogue manager to assist the creation of new entries in FRBR and an automatic procedure —schematically depicted in figure 6— to migrate the old MARC21 records into the new database. The migration process first loads

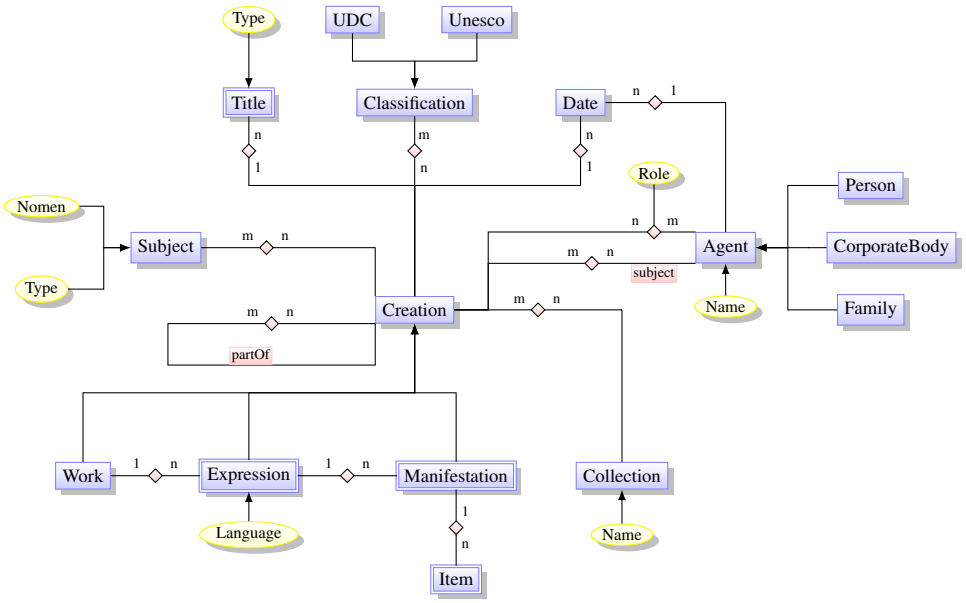

Fig. 5. Diagram of the Entity-Relationship model of the relational database.

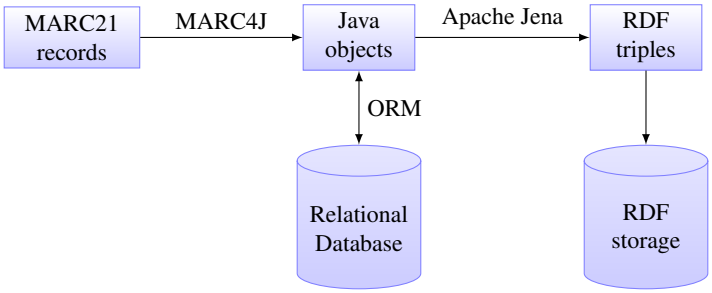

Fig. 6. Schematic representation of the migration and conversion process.

MARC21 records using the MARC4J[5] library and creates in-memory Java objects representing the FRBR entities according to the Library of Congress guidelines[6] for the MARC to FRBR transformation. Then, a transducer implements the mapping of each object field to the appropriate field in the database. The transducer uses the Hibernate ORM[7] platform which supports object/relational mapping and provides a simple interface to store and retrieve objects. ORM essentially defines the mapping between properties and columns and between classes and tables, which is declared in a configuration file.

3.2. From FRBR to RDA Linked Open Data

Once the descriptive metadata are stored in a relational database with an FRBR model, they can be published as semantic information with the method described below. The publication of RDA in RDF format first loads the fields in the database as in-memory Java objects with the Hibernate ORM library. The Apache Jena library[8] is then used to map of the object fields into the RDA components for the RDF graph (nodes and connections), as illustrated figure 8. As shown in figure 7, the core vocabulary consists of the elements and relationship designators recently approved by the Joint Steering Committee on Development of RDA and available at the RDA Registry.[9] Whenever

[5] https://github.com/marc4j
[6] http://www.loc.gov/marc/marc-functional-analysis/source/table3.pdf
[7] http://hibernate.org/orm

[8] https://jena.apache.org
[9] http://www.rdaregistry.info

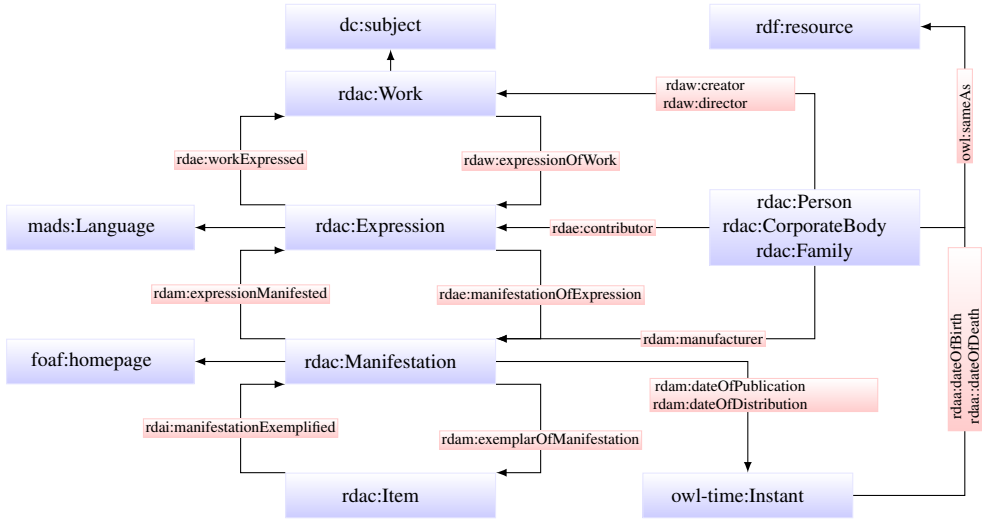

Fig. 7. The ontology (concepts and relations) describing the catalogue entries is based on the RDA, RDF, OWL, FOAF and Dublin Core vocabularies. Tag prefixes denote different name-spaces (the source ontology): RDA Class (rdac), Work (rdaw), Expression (rdae), Manifestation (rdam), Item (rdai), or Agent (rdaa); Resource Description Framework (rdf); Dublin Core (dc); Library of Congress Metadata Authority Description Schema (mads); Friend of a Friend (foaf); OWL time ontology (owl-time).

a relationship could not be described using RDA elements, then popular vocabularies were applied. For example, the OWL-Time ontology[10] has been used to describe temporal events such as publication years; external content, hosted by partner libraries, was described with FOAF elements [4] and subjects triples were created with the Dublin Core[11] property *dc:subject*. The identifiers in `data.cervantesvirtual.com` —see the RDF sample in figure 9—, are prefixed with their domain and the entity type.

The mapping between records in the database and the fields in the RDF triples exploits the Jena capacity to handle *models* (RDF graphs) and *statements* (RDF triples). A statement links a resource (the RDF subject) and a property (the RDF predicate) with a second resource or a literal text (the RDF object). A snippet of the code is shown in figure 8 and a sample output in figure 9.

The maintenance of the RDF data generated through the process described above is supported by three automatic procedures for the management of the content:

- Rebuild all RDF triples from the database.
- Incremental addition of new RDF triples.
- Data backup and restore operations.

[10] `www.w3.org/TR/owl-time`
[11] `http://dublincore.org/documents/dces`

Fully rebuilding the dataset may require a few hours but the incremental construction runs in real time and can be scheduled to take place periodically so that the published data are synchronized with the database content.

4. Results

The automatic procedure described in section 3 has been applied to transform successfully over 200000 bibliographic records and 50000 authority files. The relational model mapping was enriched with information about those relationships —for example, the *partOf* relation between *creations* and *collections* stored in MARC field 773— amenable to automatic extraction. Some manual work by the cataloguing staff was however necessary to identify relations which had not been stored in MARC records —for example, editions and translations of a single work— or connections with external datasets such as the VIAF compilation of authors.

The gateway *data.cervantesvirtual.com* —whose main features are summarized in table 1— publishes high quality linked open data —what is called *five-star open data* [10]— since:

1. the dataset is available on the web;
2. it is published as structured data;

```
public void addToModel(Model model) {
    String subject = this.getURI().toString();
    Property predicate =
       model.createProperty("http://www.w3.org/1999/02/22-rdf-syntax-ns#type");
    Resource object =
       model.createResource("http://rdaregistry.info/Elements/c/Manifestation");

    model.add(subject, predicate, object);
}
```

Fig. 8. Snippet of a Jena-based function declaring an entity to be a Manifestation for the RDF-graph.

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <rdf:Description rdf:about="http://data.cervantesvirtual.com/manifestation/242532">

    <rdf:type rdf:resource="http://rdaregistry.info/Elements/c/Manifestation"/>

    <uniformResourceLocator xmlns="http://rdaregistry.info/Elements/m/"
        rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
            http://www.cervantesvirtual.com/nd/ark:/59851/bmcwd3w6
    </uniformResourceLocator>

    <expressionManifested xmlns="http://rdaregistry.info/Elements/m/"
        rdf:resource="http://data.cervantesvirtual.com/expression/169618"/>

    <workManifested xmlns="http://rdaregistry.info/Elements/m/"
        rdf:resource="http://data.cervantesvirtual.com/work/22148"/>

    <title xmlns="http://rdaregistry.info/Elements/m/"
        rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
            El caballero de Illescas
    </title>

    <placeOfPublication xmlns="http://rdaregistry.info/Elements/m/"
        rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
            Alicante : Biblioteca Virtual Miguel de Cervantes, 2002
    </placeOfPublication>

    <dateOfDistribution xmlns="http://rdaregistry.info/Elements/m/"
        rdf:resource="http://data.cervantesvirtual.com/date/2002"/>

    <placeOfProduction xmlns="http://rdaregistry.info/Elements/m/"
        rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
            Madrid, por Juan de la Cuesta, a costa de Miguel de Syles, 1620.
    </placeOfProduction>

    <printer xmlns="http://rdaregistry.info/Elements/m/"
        rdf:resource="http://data.cervantesvirtual.com/person/5195"/>

    <publishersName xmlns="http://rdaregistry.info/Elements/m/"
       rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Siles, Miguel de
    </publishersName>

    <publisher xmlns="http://rdaregistry.info/Elements/m/"
        rdf:resource="http://data.cervantesvirtual.com/person/5257"/>

    <dateOfPublication xmlns="http://rdaregistry.info/Elements/m/"
        rdf:resource="http://data.cervantesvirtual.com/date/1620"/>

    <exemplarOfManifestation xmlns="http://rdaregistry.info/Elements/m/"
        rdf:resource="http://data.cervantesvirtual.com/item/239314"/>

  </rdf:Description>
</rdf:RDF>
```

Fig. 9. The RDF output for the novel *El caballero de Illescas*.

Table 1
Main features of the RDF *linked open data* set.

Main address	`data.cervantesvirtual.com`
Description	`.../void.ttl`
Site-map	`.../sitemap.xml`
Vocabularies	23
No. of classes	18
No. of properties	121
No. of triples	9,280,486
SPARQL access	`.../sparql`

3. data are in a standard, non-proprietary format (RDF);
4. it uses (resolvable) URIs to denote things; and
5. the content links to other datasets.

The Biblioteca Virtual Miguel de Cervantes data are published under the Creative Commons Public Domain Dedication License [12].

Even if RDA provides a comprehensive vocabulary, some issues could not be fully addressed. For example, there are different types of inclusions in a collection (such as volumes in a journal, articles in a volume, or books in a series) but RDA provides a single generic relationship *wholePartManifestationRelationship* which does not allow to specify the type of containment.

Readable and meaningful identifiers (URIs) play a crucial role in facilitating users and developers the usage of the referenced data. Our dataset therefore adheres to a number of established patterns [6]:

1. External identifiers are reused when available.
2. A readable type description of the entity is referenced within the path of the URI, for example, `language/es` or `date/2004`.

The RDF repository holds near 8000 links to different external datasets, as detailed in table 2. As customarily done, such links are described through the `owl:sameAs` relationship and they introduce the rich connectivity promoted by the *linked open data* philosophy.

Several options to provide SPARQL access to the RDF storage were evaluated, including OpenLink Virtuoso[13], 4Store[14], and Sesame[15]. The last one was selected in order to implement the access to the data, since it is an open-source Java framework which proved to be light-weight and satisfied the requirements by supporting full-text queries, batch indexing, and database transactions.[16] An open-source interface[17] was added in order to simplify the creation of queries and the visualization of results.

Table 2
Links to external data in the RDF repository.

Target	Number
`viaf.org`	5369
`isni.org`	1014
`id.loc.gov`	45
`dbpedia.org`	398
`youtube.com`	851

5. Conclusions and future work

The semantic content published in the Biblioteca Virtual Miguel de Cervantes linked open data-set includes:

- Near 9 million triples and 120 different types of relationships.
- About 195,000 FRBR *expressions*, 195,200 *manifestations* and 384,000 *items*.
- About 65,000 FRAD *persons*, 3700 *corporate bodies* and 10 *families*.
- Nearly 1200 OWL-Time *dates* of publication and distribution.
- Nearly 50 languages.

The output can be downloaded in several formats from `data.cervantesvirtual.com`: RDF, JSON, Dublin Core and ESE (Europeana Semantic Elements). An online demonstrator (`bvmcresearch.cervantesvirtual.com/arms`) has been implemented where the user can upload there a MARC21 record and obtain the RDA description in RDF format.

The *data.cervantesvirtual.com* platform is based on open-source components and implemented in the Java programming language. It has been running since February 2015 with no interruption of service.

The procedure designed in the Biblioteca Virtual Miguel de Cervantes to support FRBR, FRAD and RDA produced some additional benefits:

[12] `https://creativecommons.org/publicdomain/zero/1.0/`
[13] `http://virtuoso.openlinksw.com`
[14] `http://4store.org`
[15] `http://rdf4j.org`
[16] For an extensive comparative study of platforms, see [8].
[17] Yasgui, `http://yasgui.org`

– The FRBR hierarchical model allowed the implementation of enhanced navigation through aggregated collections such articles in journals, contributions to conference proceedings and plays in an anthology.[18]
– The relational model supports the straightforward validation of RDF data structures. This "closed-world" approach of the databases is optimal whenever the universe of information is controlled. In contrast, the "open-world" approach of RDF is more adequate when data from multiple and changing sources are merged. The combination of both technologies provides in our case a good balance between flexibility and control.

Ongoing and future work aims to create methods for the automatic extraction of implicit relationships: for example, a parser for dates and periods, or a recognizer of named persons —linking them with VIAF and DBpedia entries. We also plan to enhance the description of subjects with the creation of a thesaurus based on SKOS, a W3C recommendation for the representation of subject headings.

References

[1] American Library Ascociation, CLA, CILIP. *Anglo-American Cataloguing Rules, Second Edition*. ALA Editions, 2005.

[2] T. Baker, E. Bermès, K. Coyle, G. Dunsire, A. Isaac, P. Murray, M. Panzer, J. Schneider, R. Singer, E. Summers, W. Waites, J. Young, and M. Zeng. W3C Library Linked Data Incubator Group Final Report. Online: `http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025`, 2011.

[3] T. Baker, K. Coyle, and S. Petiya. Multi-entity models of resource description in the semantic web: A comparison of FRBR, RDA and BIBFRAME. *Library Hi Tech*, 32(4):562–582, 2014.

[4] D. Brickley and L. Miller. Foaf vocabulary specification 0.99. Online: `xmlns.com/foaf/spec`, 2014. Last visited 2015-03-01.

[5] C. Deliot. Publishing the british national bibliography as linked open data. Online: `http://www.bl.uk/bibliographic/pdfs/publishing_bnb_as_lod.pdf`, 2014.

[6] L. Dodds and I. Davis. *A pattern catalogue for modelling, publishing, and consuming Linked Data*. 2012.

[7] J. S. C. J. for Development of RDA. RDA Tool Kit: Resource Description and Access. Online: `http://www.rdatoolkit.org`, 2012.

[8] B. Haslhofer, E. M. Roochi, B. Schandl, and S. Zander. Europeana rdf store report. Technical report, University of Vienna, Vienna, March 2011.

[9] IFLA Study Group on the FRBR. *Functional Requirements for Bibliographic Records*, volume 19 of *IFLA Series on Bibliographic Control*. München: K.G. Saur Verlag, 1998.

[10] K. Janowicz, P. Hitzler, B. Adams, D. Kolas, and C. Vardeman. Five stars of linked data vocabulary use. *Semantic Web journal*, 2014.

[11] M. Lei Zeng, M. Žumer, and A. Salaba, editors. *Functional Requirements for Subject Authority Data (FRSAD) - A Conceptual Model*, volume 43 of *IFLA series on bibliographic control*. De Gruyter Saur, 2011.

[12] J. Marden, C. Li-Madeo, N. Whysel, and J. Edelstein. Linked open data for cultural heritage: evolution of an information technology. In M. J. Albers and K. Gossett, editors, *Proceedings of the 31st ACM international conference on Design of communication, Greenville, NC, USA, September 30 - October 1, 2013*, pages 107–112. ACM, 2013.

[13] R. S. Muñoz. Launching of beta version of datos.bne.es, a lod service and a frbr-based catalogue view. *Newsletter of the Standing Committee of the IFLA Cataloguing Section*, (42):13–21, December 2014.

[14] G. E. Patton, editor. *Functional Requirements for Authority Data - A Conceptual Model*, volume 34 of *IFLA series on bibliographic control*. München: K.G. Saur, 2009.

[15] A. Simon, S. Peyrard, V. Michel, and A. D. Mascio. We grew up together: data.bnf.fr from the BnF and Logilab perspectives. Online: `http://ifla2014-satdata.bnf.fr/pdf/iflalld2014_submission_Simon_DiMascio_Michel_Peyrard.pdf`, 2014.

[16] Standing Committee of the IFLA Cataloguing Section, editor. *International Standard Bibliographic Description (ISBD). Consolidated Edition*. IFLA Series on Bibliographic Control 44. De Gruyter Saur, 2011.

[17] D. Van Deursen, C. Poppe, G. Martens, E. Mannens, and R. Van de Walle. Xml to rdf conversion: a generic approach. In P. Nesi, K. Ng, and J. Delgado, editors, *Fourth International Conference on Automated Solutions for Cross Media Content and Multi-channel Distribution*, pages 138–144. IEEE Computer Society, 2008.

[18] D. Vila-Suero and A. Gómez-Pérez. datos.bne.es and marimba: an insight into library linked data. *Library Hi Tech*, 31(4):575–601, 2013.

[19] W3C World Wide Web Consortium. SPARQL Query Language for RDF. Online: `http://www.w3.org/TR/rdf-sparql-query`, 2008.

[18] See, for instance, the cross links between the hierarchical components in `www.cervantesvirtual.com/obra/anales-galdosianos--2`.