Journal Title
XX(X):1–20
©The Author(s) 2016
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Evaluating the quality of linked open data in digital libraries

Gustavo Candela[1] and Pilar Escobar[1] and Rafael C. Carrasco[1] and Manuel Marco-Such[1]

Abstract

Cultural heritage institutions have recently started to share their metadata as *linked open data* in order to disseminate and enrich them. The publication of large bibliographic datasets as linked open data is a challenge that requires the design and implementation of custom methods for the transformation, management, querying and enrichment of the data. In this report, the methodology defined by previous research for the evaluation of the quality of linked open data is analyzed and adapted to the specific case of RDF triples containing standard bibliographic information. The specified quality measures are reported in the case of four highly relevant libraries.

Keywords

Linked Data Quality, Data Quality Metrics, Linked Open Data, Digital Libraries

1 Introduction

The *semantic web* as a concept was introduced by Tim Berners-Lee in 2001 as a means to provide structure to the content of web pages.[1] The objective of the semantic web is that any entity (e.g., an individual or an organization) and any relationship between entities can be encoded on the web. *Linked Open Data* (LOD) is considered as a methodology with which to promote and facilitate the creation and reuse of semantic content. In 2010, Berners-Lee proposed 5 incremental criteria to characterize LOD. According to these criteria, LOD should be

1. available on the web with an open license;
2. available as machine-readable structured data;
3. distributed using non-proprietary formats;
4. using open standards —such as the Resource Description Framework (RDF),[2] SPARQL Protocol, and RDF Query Language (SPARQL)[3]—, and
5. linked to other repositories.

Applying the LOD concepts to the cultural heritage domain has since become an active and challenging field[4]: many galleries, libraries, archives and museums are currently exploring ways in which to convert their data into RDF and create new interfaces so as to provide a richer experience for their users.[*] The adoption of LOD maximizes metadata value, facilitates the connection of content silos with other organizations and datasets, provide a smart search context, and enable the use of synonyms and locations to enhance the discoverability and impact of culture heritage.[5,6] In addition, LOD enables the integration of the rich collections from the cultural heritage institutions into the semantic web which has become the norm for search engines in order to produce highly relevant search results.[7]

Unfortunately, the publication of bibliographic information as open data often requires intensive preprocessing, since metadata are primarily expressed in natural language. Critical choices must also be made in regards to the metadata vocabulary used to describe the objects, the ontologies employed to specify the connections between them, and the technology applied to convert the catalogue.

Several large open *knowledge bases*, —i.e., public repositories containing information that provides a wide and cross-domain coverage—, have been created in parallel, some of the most popular of which are DBpedia,[8] Wikidata[9] and YAGO.[10]

[1] Departamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante, carretera Sant Vicent s/n, 03690 Sant Vicent del Raspeig, Alicante (Spain)

Corresponding author:
Gustavo Ca.
Email: gcandela@ua.es

[*] The prefixes used to abbreviate RDF vocabularies can be found in the appendix (Table 10).

The term *knowledge graph* (KG) is often used to designate knowledge bases in the context of semantic web, although the exact definition of this term is still controversial since it has been adopted by companies and academia to describe different knowledge representation applications.[11] In the context of semantic web, a KG can be interpreted as the entire web including entities identified by links and relations according to a cross-domain ontology. The richness of the links with these open knowledge bases is clearly one of the indicators of the quality of a repository, as stated in the last dimension of the 5-star definition of LOD.

Some approaches have reused in innovative ways open data published by libraries enhancing the model of the original data and combining several datasets.[12,13] Typical uses include linking and enriching with external repositories, visualization interfaces and charts, and content analysis. However, making the choice of the best dataset is a challenge for data researchers as data quality is critical.[14,15]

The purpose of this paper is to analyze quality dimensions of LOD published by libraries, and subsequently apply these concepts to a number of cases in which the repository aims to comply with the full 5-star specification of LOD, such that their datasets are described with sufficient detail and the content becomes regularly updated. The results of this study could then be used to identify candidate datasets for reuse and enrichment.

The main contributions of this paper are the following: (a) the benchmark and the results obtained after the quality assessment; (b) the proposal of a gold standard based on RDA; and (c) the definition of a new criterion to ensure accuracy.

The paper is organized as follows: after a brief description of the state of the art in Section 2.1, Section 2.2 describes the methodology to create a benchmark of linked-data repositories. Section 3 introduces the four repositories that will serve as benchmark and discusses the methodology employed to evaluate linked data in digital libraries (DLs) and shows the results of their application. The paper concludes with an outline of the methodology adopted, general guidelines for the use of the results and future work.

2 Linked data repositories and digital libraries

2.1 Overview

The descriptive metadata of bibliographic content – which is stored as, for example, MARC records– were traditionally created and interpreted by humans. Even if those records followed specifications such as the Anglo-American Cataloguing Rules, Second Edition[16] (AACR2) and the International Standard Bibliographic Description[17] (ISBD), the textual descriptions therein could not be easily interpreted by computers, a common requirement in contemporary web-connected environments. The FRBR family of conceptual models[18] and the Resource Description and Access (RDA) specification[19] provide a modern framework for bibliographic information. However, the translation of the old records into the new format has a significant cost,[20] since libraries usually host large catalogs that must be revised manually for an accurate transformation of the data.

A growing number of cultural institutions are applying semantic web technologies and creating LOD projects. For example, the *Library of Congress Linked Data Service* (`id.loc.gov`) provides access to authority data, such as the *Library of Congress Subject Headings* and the MARC geographic areas. In 2011, the BnF published `data.bnf.fr` by aggregating information concerning authors, works, and subjects that was scattered among various catalogs. The BNE has recently migrated its databases to RDF and published them at `datos.bne.es`.[21] The BVMC catalog has also been migrated to RDF triples, which basically employ the RDA vocabulary to describe entities.[22]

Free and open knowledge bases such as Wikidata have in the mean time been growing in popularity. Wikidata allows the description of individual objects by means of properties which are proposed and defined in a participatory manner, and, if there are enough supporters and a consensus is reached, the property is eventually created by an administrator.[1]

Wikidata has raised interest in the cultural heritage domain as it offers new opportunities for the participation of the community in order to save time and energy of cultural heritage professionals. The benefits of being linked to Wikidata are (i) rich results enhanced with the information provided by KGs are becoming the standard output of search engines, and being connected to such repositories is crucial in order to increase visibility and establish a strong online presence; (ii) new routes of validation between different resources and toward better integration are opened up[23]; (iii) expertise is contributed by means of volunteers and researchers around the world who can connect the items with other collections; and (iv) Wikidata allows the execution of SPARQL federated queries in order to call out a number of external databases, including Europeana, the BVMC, and the BNE.[24]

In general, benchmarks provide an experimental basis for evaluating and comparing the performance of computer systems, information retrieval algorithms, databases, and many other technologies.[25,26] Moreover, the possibility of replicating existing results promotes further research.[27] Library benchmarks based on LOD repositories are relevant because (i) they help to compare the available repositories and to meet the needs of the consumers; (ii) researchers can address new challenges improving the methodology and including new repositories; and (iii) organizations can benefit from shared best practices when publishing their LOD repositories.

Several new approaches provide data quality criteria according to which linked data repositories can be analyzed. They have contributed to understand and specify data quality on several dimensions (e.g., accuracy, completeness, licensing).[15,28,29] These efforts are mostly concentrated on quality evaluation of KGs, which focus on general knowledge rather than specific domains such as literature. Previous work has described the adoption of linked data by libraries, archives, and museums, identifying the current trends and challenges.[30,31] The specific vocabularies used in the LOD repositories published by libraries allow for greater expressiveness since they are addressed to the bibliographic content. For instance, the use of different roles, such as editor and illustrator, when assigning an author to a work. To the best of our knowledge, none has been carried out to perform a quantitative evaluation of the linked open data published by DLs.

This paper is based on the data-quality criteria for the KGs previously published[15] which have been analyzed here and adapted to the context of DLs. They have been then applied to evaluate the linked data published by four relevant libraries: the Biblioteca Nacional de España (BNE), the Bibliothèque nationale de France (BnF), the British National Bibliography (BNB) and the Biblioteca Virtual Miguel de Cervantes (BVMC). The results could be used to identify the most appropriate library for a specific purpose by weighting the scores obtained for every quality criterion.

2.2 Methodology for the selection of repositories

The main goal of this study is to provide the linked data community with a benchmark for the comparison and evaluation of data quality in digital libraries. Since the number of libraries publishing linked data has grown rapidly, identifying subjects –candidates for the assessment– is a critical factor for the success of a benchmark. Other approaches propose methodologies to identify subjects that consider various attributes ranging from technical issues to cultural aspects.[26,32]

In this approach, the subject repositories in the benchmark must meet the following criteria: accessible under an open license; a public SPARQL endpoint available; the content becomes regularly updated; and a public web interface available.

Suitable subject datasets can be identified in public repositories such as Wikidata and LOD Cloud,[2] and also in journal articles addressing DLs. However, some of the items can be out of date, may lack uniform structure or use invalid URLs.

The set of subjects can be further refined through the analysis of additional characteristics such as: the number of vocabularies used; the number of publications; the number of Wikidata properties; being described by vocabularies based on, or derived from, FRBR; and the number of awards or citations received. The number of awards and scientific publications generated by a DL can be retrieved by exploring their websites as well as repositories of scientific communications such as Scopus, DBLP and Google Scholar.

The list of potential subjects can be evaluated with a variety of techniques based on multi-attribute decision-making tools. For example, the alternatives to alternatives scorecard uses a matrix in which columns are labelled with subjects, rows are labelled with criteria, cells contains a numerical performance measure, and the best subject for each attribute is then highlighted. Another popular and visual technique are *polar charts*, where rays are drawn from the centre of a circle –each one associated to an attribute with length proportional to the rating– and the subject covering the larger area is considered the best choice.

3 Assessing the data-quality of LOD in digital libraries

This section introduces the four repositories that will serve as benchmark and the results obtained by applying the procedure to evaluate each criterion proposed in Table 3.

3.1 A benchmark of linked-data repositories

In order to find suitable subject datasets, we have applied the methodology described in Section 2.2. We identified datasets in the current LOD Cloud whose description contains terms such as *library* or are included in Section 2.1. Some subjects were removed because of being out of date or using not valid URLs. Table 1 presents a preliminary list of candidates.

We then used polar charts to identify which LOD repositories are most suitable for the study. Every axis on the polar chart corresponds to one of the following features: vocabularies; publications; Wikidata properties; FRBR; and prizes. The axis values have been normalized and the global score is computed as the area of the polar chart –as shown in Figure 3 for the BnF. If the subject does not provide a SPARQL endpoint, the area is not computed.

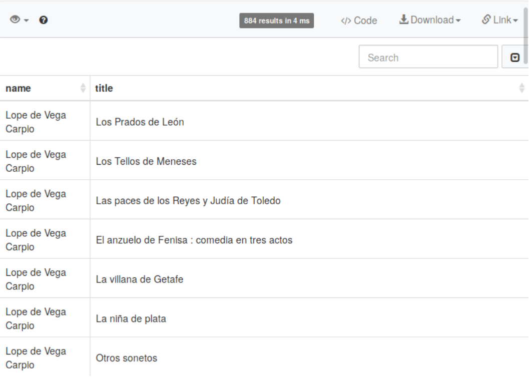
As a result of the evaluation, four libraries (BnF, BNE, BNB and BVMC) were selected which implement the LOD concepts. Although the number

```
PREFIX rdaa: <http://rdaregistry.info/Elements/a/>
SELECT ?name ?title
WHERE {
    wd:Q165257 wdt:P2799 ?id .
    wd:Q165257 wdt:P1559 ?name .
    BIND(
        uri(concat("http://data.cervantesvirtual.com/person/", ?id))
        AS ?bvmcID
    )
    SERVICE<http://data.cervantesvirtual.com/openrdf-sesame/
            repositories/data> {
       ?bvmcID rdaa:authorOf ?work .
       ?work rdfs:label ?title
    }
}
```

Figure 1. A SPARQL query retrieving the works of Wikidata author `wdt:P2799` (Lope de Vega) from a remote repository —that specified after the `SERVICE` keyword. The output is shown in Figure 2.

Table 1. Criteria for the selection of the subjects in which the global score is the area.

Subject	License	SPARQL endpoint	Web interface	Maturity	Update	Vocabularies	Publications	Wikidata properties	FRBR based	Prizes	Area
BnF	Open Licence	1	1	1	1	10	6	1	1	1	7.275
Europeana	CCO	1	1	1	1	5	1	1	0	0	0.085
BNB	CCO	1	1	1	1	11	2	0	0	0	0.329
BNE	CC0	1	1	1	1	12	2	1	1	0	0.969
LOC	Public domain	0	1	1	1	5	1	3	1	0	-
BVMC	Public domain	1	1	1	1	14	2	3	1	2	6.975
Deutsche Nationalbibliothek (DNB)	CC0	0	1	1	1	-	0	1	1	0	-
National Széchényi Library (NSZL)	Other (Open)	0	0	1	0	4	0	1	0	0	-
National Library of Greece Authority Records (NLG)	Other (Open)	1	0	1	1	8	0	1	0	0	0



name	title
Lope de Vega Carpio	Los Prados de León
Lope de Vega Carpio	Los Tellos de Meneses
Lope de Vega Carpio	Las paces de los Reyes y Judía de Toledo
Lope de Vega Carpio	El anzuelo de Fenisa : comedia en tres actos
Lope de Vega Carpio	La villana de Getafe
Lope de Vega Carpio	La niña de plata
Lope de Vega Carpio	Otros sonetos

Figure 2. Output of the SPARQL query in Fig. 1.

of triples varies considerably among the datasets, these libraries mainly publish information about works, authors and subjects –see Figure 4 for the fraction of entities in each FRBR group.

The main features of the selected repositories are:

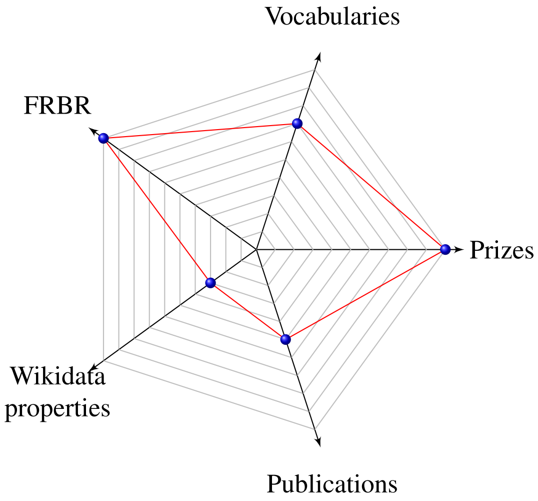


Figure 3. Polar chart that shows the area according to the values for the BnF in Table 1.

1. `datos.bne.es`, the linked data service of the Biblioteca Nacional de España. The dataset is the

result of an experiment that was developed jointly by the BNE and the Ontology Engineering Group from the Universidad Politécnica de Madrid. The metadata have been transformed into models, structures and vocabularies following the FRBR architecture proposed by the International Federation of Library Associations and Institutions (IFLA), thus making them more interoperable and reusable. Traditional MARC21 files were processed with Marimba, a tool developed by the research group to map subfields onto properties. Marimba also supports the enrichment of data with external resources, such as VIAF and Wikipedia.

The BNE collection contains 2 million works, 1.4 million expressions, one million manifestations, and 1.4 million items. Almost 1.5 million authors are represented by *person* and *corporate body* classes and 0.65 million subjects are described using `skos:Concept`.

2. `data.bnf.fr`, published in 2011 by the Bibliothèque nationale de France, which aggregates information concerning authors, works, and subjects that were formerly scattered among various catalogs. These data are published in RDF using a vocabulary based on the FRBR model in which objects are referenced through the use of ARK (Archival Resource Key) identifiers. The information is stored in different formats, including RDF, JSON, and HTML.[33] The platform is based on CubicWeb,[3] an open-source platform used to develop semantic web applications.

   The BnF repository contains about 21 million entities in FRBR group 1 (0.65 million works, 10 million expressions and 10 million manifestations) in RDA vocabulary.[34] Moreover, approximately 2 million authors are also described by means of the *foaf:Person* class and 0.6 million subject headings are linked to RAMEAU entries.[35]
3. `bnb.data.bl.uk`, the British National Bibliography linked data platform which supports the SPARQL query language and delivers RDF and JSON output. The dataset has been modeled upon RDF vocabularies, such as Dublin Core, the Bibliographic Ontology (BIBO), and Friend of a Friend (FOAF). The full dataset is available for download.[36]

   The BNB repository contains 2 million authors represented as `foaf:Agent` entities and 1.5 million subjects linked to Library of Congress Subject Headings.
4. `data.cervantesvirtual.com`, the Biblioteca Virtual Miguel de Cervantes open-data repository. The 200,000 entries in the catalog were transformed into RDF triples by employing principally the RDA vocabulary.[22]

   The BVMC dataset describes 0.2 million bibliographic records and 0.1 million authors based on the RDA vocabulary.[37]

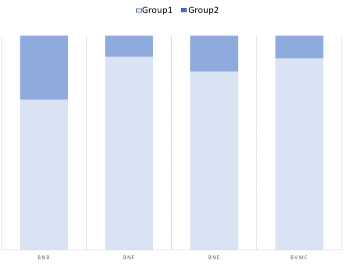


Figure 4. Distribution of entities by FRBR group: group 1 includes works, expressions and manifestations; group 2 includes persons, corporate bodies and families.

3.2 Data quality analysis

The data-quality criteria to evaluate datasets in the LOD context, and KGs in particular, listed in Table 3 employ the concepts of criteria, dimensions and categories originally proposed by Wang and Strong[38] in the context of data quality.[15]

A data-quality criterion is a function with values in the range of 0–1 that scores a particular feature – such as the syntactic validity of literals. A data-quality dimension comprises one or more criteria which are, in turn, grouped into categories.

The dimensions and criteria listed in Table 3 are defined for KGs. Libraries use, however, specific vocabularies for the description of their resources which include rich and expressive relationships between editions of the same work, the specification of a sequence between works (e.g., continuation of), or the use of multiple roles (e.g., illustrator and editor) when assigning an author to a work.

We have adapted the procedure to evaluate each criterion proposed listed in Table 3 to the specificities of bibliographic content, detailed in Sections 3.3–3.13. Only one additional criterion has been introduced — namely, duplicate entries, see the last item in Section 3.3—, which measures the degree of redundancy in the entries of the repository. The analysis below and the figures in Table 9 were obtained in the period July–November 2018.

3.3 Accuracy

Definition. According to the literature,[38] the accuracy dimension determines *the extent to which data are correct, reliable, and certified free of error*.

Table 2. Key figures as regards the benchmark repositories.

	BNB	BnF	BNE	BVMC
Number of triples	151,779,391	334,457,101	79,448,899	13,155,339
Number of entities	25,789,090	33,804,333	7,860,809	1,499,362
Number of classes	30	32	28	33
Number of properties	75	94	286	182

Table 3. The data-quality criteria classified by category and dimension.

Category	Dimension	Criterion
Intrinsic category	Accuracy	Syntactic validity of RDF documents Syntactic validity of literals Syntactic validity of triples
	Trustworthiness	Trustworthiness on KG level Trustworthiness on statement level Using unknown and empty values
	Consistency	Check of schema restrictions during insertion of new statements Consistency of statements w.r.t. class constraints Consistency of statements w.r.t. relation constraints
Contextual category	Relevancy	Creating a ranking of statements
	Completeness	Schema completeness Column completeness Population completeness
	Timeliness	Timeliness frequency of the KG Specification of the validity period of statements Specification of the modification date of statements
Representational data-quality	Ease of understanding	Description of resources Labels in multiple languages Understandable RDF serialization Self-describing URIs
	Interoperability	Avoiding blank nodes and RDF reification Provisioning of several serialization formats Using external vocabulary Interoperability of proprietary vocabulary
Accessibility category	Accessibility	Dereferencing possibility of resources Availability of the KG Provisioning of public SPARQL endpoint Provisioning of an RDF export Support of content negotiation Linking HTML sites to RDF serializations Provisioning of KG metadata
	License	Provisioning machine-readable licensing information
	Interlinking	Interlinking via owl:sameAs Validity of external URIs

In the context of libraries, accuracy is a critical indicator of quality since users expect accurate and error-free data.[39] Traditional issues in DLs are metadata and typographical errors, the size of the collections, and the complexity of the new formats may lead to duplicated entities and syntax errors when producing the documents.

Assessment. The accuracy dimension was evaluated by means of three criteria listed in Table 3. In order to assess the criteria, a list of all authors was retrieved from their SPARQL endpoints, and a random sample of 100 authors was selected per each DL.[4] The criteria in Table 3 are complemented with the automatic detection of duplicated entities:

- *Syntactic validity of RDF documents*. The use of standard tools and software is recommended when creating RDF documents. Syntax errors in RDF can be identified using tools such as the W3C RDF Validator.[40] The criterion was originally defined as:

$$m_{\text{synRDF}} = \begin{cases} 1 & \text{if all RDF documents are valid} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The W3C RDF validator was used to assess the RDF documents of the random sample, and it was found that all of them provide syntactically valid RDF documents.

- *Syntactic validity of literals*. The literal values stored in the DLs can be used for this purpose by means of regular expressions. Syntactic rules are patterns to test dates and identifiers in DLs. The RDF graph G consists of RDF triples (s, p, o) and a set of literals, L. The original methodology defines:

  $$m_{\text{synLit}} = \frac{|\{G \wedge L \wedge o \,\text{is valid}\}|}{|\{G \wedge L\}|} \qquad (2)$$

  Common properties such as dates associated with authors, International Standard Name Identifiers (ISNI) and International Standard Serial Numbers (ISSN) were tested against their patterns.
  There are 0.3 million `bibo:issn` triples in the BnF. In the BNB, 179 out of 0.1 million `bibo:issn` were not syntactically correct.[†] In the BVMC, a single ISBD property[41] is used to store the information regarding ISSN and ISBN, thus hindering automatic validation. All triples in the BVMC (about ten thousand) were found to be correct. Although some works in the BNE contained an ISSN, they were not available in RDF format at the time of this analysis.
  The sample of 100 authors per library was tested using a semi-automatic process. First, all properties were gathered and processed automatically. Then, a manual revision was performed in order to identify inconsistencies. All the literal values verified using the relation *date of birth* were syntactically correct. The list of the RDF-compatible types specifies that the type `xsd:date` must be encoded in the yyyy-mm-dd format (with or without a timezone).[5] Some dates were, however, found to include qualifiers —such as b., d., ca., fl., ?, and cent.[‡]
  The ISNI is a code with which to uniquely identify public identities of contributors to media content, such as books and articles. Each identifier is a 16 digit number which can also be displayed as four blocks with four digits in each block. A sample of 500 ISNIs was selected per library by accessing their SPARQL endpoints and all of them were found to be correct.
- *Semantic validity of triples*. The semantic validity of triples is evaluated with a reference dataset that serves as a gold standard S. The criterion measures the extent to which the triples in the repository G and in the gold standard S have the same values. Then we can state:

  $$m_{\text{semTriple}} = \frac{|G \wedge S|}{|G|} \qquad (3)$$

  The random sample of 100 authors was compared with entries in the Virtual International Authority File (VIAF), a service that integrates access to major authority files.[42] Dates of birth, places of birth, dates of death, places of death and alternate names, when available, were retrieved from VIAF and manually checked against the values in the sample. The triples in all samples were found to be correct.
- *Duplicate entities*. One method that can be employed to recognize multiple identifiers for a single entity in a repository is that of inspecting the links from external knowledge bases. For example, the authors *Polo, Marco, 1254-1324* and *Marco Polo* in BVMC are both identified by Wikidata as `wd:Q6101`.
  A score can, therefore, be computed as the rate of links in Wikidata with a duplicate target (for example, if a Wikidata entry is linked to 3 instances in the repository, 2 are duplicates). Formally, let n_w^u be the number of unique entities linked to Wikidata, and n_w the number of links to Wikidata, then:

  $$m_{\text{checkDup}} = \frac{n_w^u}{n_w} \qquad (4)$$

  The amount of duplicate entries can be obtained with a query like that shown in Fig. 5 and the results are depicted in Table 4.

```
SELECT ?s (COUNT(?id) AS ?total)
WHERE { ?s wdt:P268 ?id }
GROUP BY ?s
HAVING (COUNT(?id) > 1)
```

Figure 5. SPARQL query retrieving duplicate identifiers in the BnF. Wikidata property *wdt:P268* is `BnF Id`.

Discussion. All the datasets evaluated attain a high score in the *accuracy* dimension, remarkably the BnF. Some specific features of this type of repositories, –such as providing a year rather than full dates– were identified. A new criterion has been introduced in this dimension that evaluates the number of duplicate

[†] Such as an ISSN with ten digits for the item http://bnb.data.bl.uk/id/series/Developmentsinfoodscience0444416889, requested on October 1, 2018

[‡] For example, the date of birth of *Gonzalo de la Cerda* is encoded at the BnF as `fl. 15--`.

Table 4. Number of duplicate entities per library.

Wikidata property	no. of links	no. of duplicates
BnF ID (P268)	447,453	2042 (0.46%)
BNE ID (P950)	139,023	821 (0.59%)
BNE journal ID (P2768)	259	1 (0.39%)
BVMC person ID (P2799)	10766	356 (3.31%)
BVMC work ID (P3976)	512	5 (0.98%)
BVMC place id (P4098)	20	0 (0.00%)
BNB person ID (P5361)	32,745	647 (1.98%)

Table 5. Possible scores according to the criterion trustworthiness on dataset level.

Description	Score
Manual data curation, manual data insertion in a closed system	1
Manual data curation and insertion, both by a community	0.75
Automated data curation, data insertion by automated knowledge extraction from structured data sources	0.25
Automated data curation, data insertion by automated knowledge extraction from unstructured data sources	0

entities. As the number of duplicates is not large, they could easily be revised for greater accuracy. At the time of writing this report, no property in Wikidata was linked to the BNB, and a new property identifying people in the BNB was, therefore, suggested by the authors.[6]

3.4 Trustworthiness

Definition. Trustworthiness is defined as *the degree to which the information is accepted to be correct, true, real, and credible*.[43]

Assessment. Trustworthiness is evaluated at three levels:

- *Trustworthiness on dataset level*. The criterion is originally defined as shown in Table 5.
  All the libraries perform automatic conversions to LOD,[21,22,33,36] which corresponds to the score 0.25 in Table 5.
- *Trustworthiness on statement level*. The fulfillment of this criterion means that a provenance vocabulary is used to describe derived data. Information concerning the provenance of the data can be encoded, for example, by using the `prov:wasDerivedFrom` property in the W3C-PROV ontology[44] or the `dcterms:provenance` and `dcterms:source` properties in Dublin Core. The original criterion distinguishes between provenance information for triples and provenance information for resources and it is defined as:

$$m_{\text{fact}} = \begin{cases} 1 & \text{provenance on statement level is used} \\ 0.5 & \text{provenance on resource level is used} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

  None of the libraries employed either an external or a proprietary vocabulary to store provenance information.
- *Using unknown and empty values*. Trustworthiness can be increased by supporting unknown and empty values. These statements —such as the authors of anonymous books retrieved by the query shown in Fig. 6— require unknown and empty values to be encoded with a different identifier. The criterion was originally defined as:

$$m_{\text{NoVal}} = \begin{cases} 1 & \text{unknown and empty values are used} \\ 0.5 & \text{either unknown or empty values are used} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

  None of the libraries was found to differentiate unknown values from empty records.

```
SELECT *
WHERE { ?works wdt:P50 wd:Q4233718 }
```

Figure 6. SPARQL query retrieving works with unknown authors. Tag `wdt:P50` represents the main creator of a written work and `wd:Q4233718` is an anonymous entity in Wikidata.

Discussion. *Trustworthiness* is not very high at this type of repositories because data are automatically extracted from supervised and structured data sources and they are not revised after creation. This criterion should probably be redefined in this context, as it was created to analyze other repositories in which the data are not curated before their publication. It would be desirable that DLs include provenance information as part of the metadata.

3.5 Consistency

Definition. Consistency is defined as *two or more values that do not conflict with each other.*[45] *Semantic consistency is the* extent to which the collections use the same values (vocabulary control) and elements for conveying the same concepts and meanings throughout.[46]

The use of controlled vocabularies facilitates consistency in DLs. However, the use of different providers and the structural complexities of OWL when representing knowledge can lead to inconsistencies. In this context, OWL allows the introduction of restrictions with regard to classes and relations in order to ensure consistency.

Assessment. Three aspects of consistency are measured:

- *Consistency of schema restrictions during insertion of new statements*. Checking the schema restrictions during the insertion of new statements are often done on the user interface in order to avoid inconsistencies. For instance, that the entity to be added has a valid entity type, as expressed by the `rdf:type` property.

$$m_{\text{checkRestr}} = \begin{cases} 1 & \text{schema restrictions are checked} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

  The user interfaces were examined and none was found to test schema constraints.
- *Consistency of statements with respect to class constraints*. This metric measures the extent to which the instance data is consistent with regard to the class restrictions. Following other approaches,[15] we limit ourselves to the class constraint `owl:disjointWith`.
  Let CC be the set of all class constraints, defined as $CC = \{(c_1, c_2) | (c_1, \text{owl:disjointWith}, c_2) \epsilon g\}$. Then, let $c_g(e)$ be the set of all classes of instance e in g, defined as $c_g(e) = \{c | (e, \text{rdf:type}, c) \epsilon g\}$. Then we can state:

$$m_{\text{conClass}} = \frac{|\{(c_1, c_2) \epsilon CC | \neg \exists e : (c_1 \epsilon c_g(e) \wedge c_2 \epsilon c_g(e))\}|}{|\{(c_1, c_2) \epsilon CC\}|} \tag{8}$$

  The definition of the vocabularies and the constraints used were revised, in the attempt to discover statements such as `owl:disjointWith`. When no information was available, the SPARQL endpoint was queried and restrictions, such as a person not also being an organization were checked —see Fig. 7. Only the BnF defines seven class constraints using the FOAF and SKOS vocabularies as a basis, and all of their triples satisfy the constraints. At least one entity in the BNE was described as both `Person` and `Corporate Body`.[§]

```
SELECT ?entity
WHERE {
     ?entity rdf:type bneonto:C1005 .
     ?entity rdf:type bneonto:C1006
}
```

Figure 7. SPARQL query retrieving resources typed simultaneously as Person (class `C1005`) and Corporate Body (class `C1006`).

- *Consistency of statements with respect to relation constraints*. This metric measures the extent to which the instance data is consistent with the relation restrictions. We evaluate this criterion by averaging over the scores obtained from single metrics $m_{conRelat,i}$ indicating the consistency of statements with regard to the relation constraints `rdfs:range` and `owl:FunctionalProperty`:

$$m_{\text{conRelat}} = \frac{1}{n} \sum_{i=1}^{n} m_{conRelat,i}(g) \tag{9}$$

  The relation `rdfs:range` specifies the type of entities that can occur at the third position in a triple and the consistency of the statements with this constraint can be checked using the SPARQL query shown in Fig. 8). In the BNE dataset, the relation `bneonto:OP1005` (*is created by*) requires an entity type `bneonto:C1006` (*Corporate Body*), but the entity type `bneonto:C1001` (*Work*) appears instead in about 2% of these

[§] The item http://datos.bne.es/resource/, requested on Nov 5, 2018

```
SELECT (COUNT(?x) as ?total)
       ?rangeType
WHERE { ?x bneonto:OP1005 ?o .
        ?o a ?rangeType }
GROUP BY ?rangeType
```

Figure 8. SPARQL query checking that the object of all the `OP1005` (*is created by*) properties in the BNE has the right type (in this case, *Corporate Body*).

relations –see Fig. 8. No issues were found for the BnF, BVMC and BNB.

Discussion. The *consistency* of data is high but schema restrictions are not checked during the insertion of new statements. This criterion may not be applicable to the evaluation of the LOD created by libraries, because the collection of data by external contributors is not currently among their objectives.

3.6 Relevancy

Definition. Relevancy is *the extent to which data is useful for the action performed.*[47]

Assessment. There is only one criterion in the relevancy dimension:

- *Creating a ranking of statements*. It is evaluated whether the DL supports a ranking of statements in order to express the relative relevance of statements.

$$m_{\text{Ranking}} = \begin{cases} 1 & \text{ranking of statements supported} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

  None of the libraries supports the ranking of statements, entities or relations, which could be used in this context to, for example, store the order of authors in the original publication.

Discussion. The *relevancy* dimension has not been considered by the libraries in our sample, as they do not provide rankings of statements, entities or relations in order to, for example, store the order of authors in the original publication.

3.7 Completeness

Definition. Completeness is *the extent to which data are of sufficient breadth, depth, and scope for the task at hand.*[38]

DLs distribute their own content, which may not cover all themes, writers or dates.

Assessment.

Table 6. RDA classes and properties used to evaluate the completeness criteria.

Class	Properties
Person	name, date of birth, date of death
Corporate body	name
Family	name, founding year
Work	form of work, title, creator
Expression	language, editor, translator
Manifestation	date of publication, note

The completeness dimension is inspected at three levels:

- *Schema completeness*. This criterion measures the extent to which classes and relations are not missing. We used a gold standard which includes entities and properties traditionally found in DLs such as person, work, name and title, based on the RDA vocabulary[7] – see Table 6.
  The schema completeness $m_{cSchema}$ is defined as the ratio of the number of classes and relations of the gold standard existing in g, no_{clatg}, and the number of classes and relations in the gold standard, no_{clat}:

$$m_{\text{cSchema}} = \frac{no_{clatg}}{no_{clat}} \tag{11}$$

  The BVMC obtains a high score when using this measure because its main vocabulary is based on RDA. The BNB is, however, based on BIBO, in which publication entities are not described as FRBR. The BNE does not provide entities typed as Family and the BnF Agent entities are based on FOAF.
- *Column completeness*. Columns completeness is defined as the rate of instances that have a specific property defined, averaged for all the properties in Table 6.
  Let H be the set of all combinations of the considered classes and relations, column completeness was originally defined as the ratio of the number of instances of class k and a relation r, no_{kr}, to the number of all instances typed as k, no_k.

$$m_{\text{cCol}} = \frac{1}{H} \sum_{k,r \epsilon H} \frac{no_{kr}}{no_k} \tag{12}$$

  The score obtained by the BNB is low because its data model is based on BIBO (Book class) and Dublin Core (creator and contributor roles), while our gold standard includes entities from the FRBR model. In the BNE, the *Family* class is absent and translators of a work are labeled with

```
SELECT DISTINCT ?writer
WHERE {
    ?writer wdt:P31 wd:Q5 .
    ?writer wdt:P106 wd:Q49757 .
    ?writer wdt:P106 wd:Q214917 .
}
```

Figure 9. SPARQL query retrieving poetry writers —in which the entity *is instance of* (`P31`) is *human* (`Q5`), *has as occupation* (`P106`) is both *person who writes and publishes poetry* (`wd:Q49757`) and *playwright*.

the generic property *participant* rather than the more specific *translator*.[¶]
No property in the BnF describes the form of a work, although this is sometimes implicit in types such as `bibo:Periodical` or `dcmitype:InteractiveResource`.[‖] The property *éditeur scientifique* in namespace *bnfroles* was taken to be equivalent to the property *editor* in our table. Some family entities in the BnF dataset are not typed as *Family*.[**]

- *Population completeness*. This criterion determines the extent to which the DL covers a basic population. Let E_s be the set of entities in the gold standard, and E_g the set of entities in g, we can define:

$$m_{\text{cPop}} = \frac{|E_s \wedge E_g|}{|E_s|} \qquad (13)$$

The coverage of entities was compared with that of Wikidata, and particularly a list of writers creating poetry and theater —see Fig. 9.

Discussion. The *completeness* dimension has two different types of criteria: all the libraries score high in the usage of the elements defined in the schema (a natural result, to a certain extent, as the schema has been fitted to their purposes) and they score low in data population because they provide only curated data based on the content of their own collections of bibliographic records and they do not have universal coverage as a principal target.

3.8 Timeliness

Definition. Timeliness of a digital object is *the extent to which it is sufficiently up-to-date for the task at hand.*[48]

Timeliness measures if the resource includes metadata about when was created, stored, accessed or cited. Users expect updated objects and time of the last freshening is a relevant quality indicator.[49]

Assessment. The timeliness dimension involves the frequency and information of the updates:

- *Timeliness frequency*. This criterion indicates how often the DL is updated. The original methodology differentiates between continuous and discrete updates.

$$m_{\text{Freq}} = \begin{cases} 1 & \text{continuous updates} \\ 0.5 & \text{discrete periodic updates} \\ 0.25 & \text{discrete non-periodic updates} \\ 0 & \text{otherwise} \end{cases} \qquad (14)$$

The frequency of updates was consulted in all the repositories.[††] When this was not available, properties such as `dcterms:created` were examined and, after which the VoID files were inspected. All libraries update the dataset more that once per year, which corresponds to a score of 0.5 in Färber's methodology. None of them provides a list of the versions with dates of publications, in contrast to repositories such as DBpedia.

- *Specification of the validity period of statements*. This criterion measures whether the repository supports the specification of starting and end dates of statements.

$$m_{\text{Validity}} = \begin{cases} 1 & \text{specification of validity period supported} \\ 0 & \text{otherwise} \end{cases} \qquad (15)$$

None of the libraries use properties –such as Wikidata *end time* (`P582`)– to specify validity, probably because bibliographic records are created as persistent objects.

- *Specification of the modification date of statements*. This criterion measures the use of dates as the point in time of the last verification of a statement represented by means of the properties `schema:dateModified` and `dcterms:modified`.

[¶] See, for example, http://datos.bne.es/edicion/Mimo0001709479.html.

[‖] See, for example, http://data.bnf.fr/ark:/12148/cb326801160#about

[**] See, for example, https://data.bnf.fr/fr/11978989/curie/

[††] See, for example, https://data.bnf.fr/en/about

$$m_{\text{Change}} = \begin{cases} 1 & \text{specification of modification dates for statements supported} \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

Modification dates are specified only in the BnF —by means of the `dcterms:modified` property. No usage of property `schema:dateModified` was found.

Discussion. The results were, in general, low for *timeliness*, with the exception of the BnF, the only case in which the modification date of statements is provided.

3.9 Ease of understanding

Definition.

The ease of understanding is *the degree to which data are understood, readable and clear.*[38]

In the context of a DL, this is focused on users and addresses issues such as using textual descriptions and descriptive URIs. Since most of libraries are local or national, they often provide their content in a single language.

Assessment. The ease of understanding is measured by means of four criteria:

- *Description of resources*. Repositories based on semantic web principles may use basic properties (for instance, `rdfs:label` and `rdfs:comment`) to describe resources. Formally, let P_{lDesc} be the set of relations that contains a label or description and U_g^{local} the set of all URIs in g with local namespace:

$$m_{\text{Descr}}(g) = |\{u|u\epsilon U_g^{local} \wedge \exists(u,p,o)\epsilon g : p\epsilon P_{IDesc}\}/\{u|u\epsilon U_g^{local}\}| \tag{17}$$

The rate of entities described with the property `rdfs:label` has been computed and found to be high in all cases.

- *Labels in multiple languages*. This criterion measures whether labels in additional languages are provided.

$$m_{\text{Lang}} = \begin{cases} 1 & \text{Labels provided in at least one additional language} \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

The textual value of a property can be encoded in multiple languages by adding attributes such as `@es`, `@fr`, etc. The BnF declares the language of the `dcterms:title` and `dcterms:description` properties, in which references to 12 languages were found. The BNE, BVMC and BNB do not include this type of information although they have some content in foreign languages.[‡‡]

- *Understandable RDF serialization*. This criterion measures the use of alternative encodings that are more understandable for humans than RDF, such as N-Triples, N3 and Turtle.[50]

$$m_{\text{uSer}} = \begin{cases} 1 & \text{Other RDF serializations than RDF/XML available} \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

The BNB and BnF provide N-Triples and Turtle serializations. The BNE disseminates only Turtle. The BVMC publishes RDF/XML and JSON-LD on the website, and additional formats can be obtained through the use of the SPARQL endpoint.[8]

- *Self-describing URIs*. Self-descriptive URIs contain a readable description of the entity rather than identifiers and they help users to understand the resource.

$$m_{\text{uURI}} = \begin{cases} 1 & \text{self-describing URIs always used} \\ 0.5 & \text{self-describing URIs partly used} \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

The BnF uses URIs with the full name of the resource. The BVMC and BNE URIs contain a readable description of the entity class and an identifier of the resource. The BNB relies on opaque URIs.

Discussion. The scores measuring *ease of understanding* are diverse, depending on the criterion: for example, only the BnF provides labels in multiple languages while the BNB does not employ self-describing URIs. No library includes both the entity and the label

[‡‡] For example, http://datos.bne.es/persona/XX1718747.html and http://bnb.data.bl.uk/doc/resource/009648286.

in the URI, which would, from our point of view, be the optimal choice for users.

3.10 Interoperability

Definition.

The interoperability enables machines to exchange information, data and knowledge in a meaningful way.[51]

Interoperability is crucial to facilitate the sharing and reuse of LOD. Providing machine readable metadata is a key aspect.

Assessment.

Interoperability involves four criteria:

- *Avoiding blank nodes and RDF reification.* This criterion tests the use of blank nodes and RDF reification.

$$m_{\text{Reif}} = \begin{cases} 1 & \text{no blank nodes and no RDF reification} \\ 0.5 & \text{either blank nodes or RDF reification} \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

  The RDF reification vocabulary — the `rdf:Statement` class and the `rdf:subject`, `rdf:predicate`, and `rdf:object` properties— are not used by the libraries. Blank nodes were also checked with the `isBlank` SPARQL operator.
- *Provisioning of several serialization formats.* This criterion measures the support of additional formats to RDF/XML for URI dereferencing.

$$m_{\text{iSerial}} = \begin{cases} 1 & \text{RDF/XML and further formats are supported} \\ 0.5 & \text{only RDF/XML is supported} \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

  All of the libraries provide results in at least RDF/XML, JSON-LD and Turtle which corresponds to score 1.
- *Using external vocabulary*. This score was obtained as the fraction of triples using an external vocabulary in their predicate.

$$m_{\text{extVoc}} = \frac{|\{(s,p,o)\epsilon g \wedge p\epsilon P_g^{external}\}|}{|\{(s,p,o)\epsilon g\}|} \tag{23}$$

  The BNB employs 67 relations from 9 external vocabularies, the BVMC 158 properties from 11 vocabularies (mainly RDA), the BNE 38 properties (in RDF, RDFS, and OWL), while the BnF uses 100 relations from 10 external vocabularies.
- *Interoperability of proprietary vocabulary*. This criterion computes the fraction of classes and relations with at least one equivalence link to classes and relations in external data sources. Equivalences can be declared by means of `owl:sameAs`, `owl:equivalentClass`, `rdfs:subPropertyOf` or `rdfs:subClassOf`.

  Let $P_{eq} = \{$owl:sameAs, owl:equivalenClass, rdfs:subPropertyOf, rdfs:subClassOf$\}$ and U_g^{ext} consists of all URIs in U_g which are external to the DL g, we can state:

$$m_{\text{propVoc}} = \{(x,p,o)\epsilon g \wedge (p\epsilon P_{eq} \wedge o\epsilon U_g^{ext})\} \tag{24}$$

  The BNE declares a high number of equivalences through the use of the `rdfs:subClassOf` property,[9] and about 62% of them link to external vocabularies. In the BnF, only one proprietary class, *Online exhibition*, is linked — to `foaf:Document`. The BnF relations are linked to FOAF, DC and RDA with a coverage of 85.3%. In the BVMC, all the classes and properties are taken from external vocabularies based mainly on RDA, FOAF, schema.org and SKOS. With regard to BNB, 35.7% of the properties are linked to external classes — in SKOS and Event[10]— by means of the `rdfs:subClassOf` relation.

Discussion. *Interoperability* is high for all repositories, as they provide a number of output formats and employ relevant external vocabularies.

3.11 Accessibility

Definition.

Accessibility is *the extent to which data are available or easily and quickly retrievable.*[38]

Accessibility requires the data to be available through SPARQL endpoints and RDF dumps. SPARQL endpoints also allow the execution of federated queries accross different datasets, enhancing and increasing the visibility of the LOD.

Assessment.

The accessibility involves a variety of criteria:

- *Dereferencing possibility of resources*. Dereferencing of resources is based on URIs that are resolvable by means of HTTP requests, returning useful and valid information. The dereferencing of resources is successful when an RDF document is returned and the HTTP status code is 200. This criterion assesses for a set of URIs whether dereferencing of resources is successful. Let U_g be a set of URIs, we can state:

$$m_{\text{Deref}} = \frac{|\text{dereferencable}(U_g)|}{|U_g|} \tag{25}$$

A random choice of 5,000 URIs was requested for all the libraries from their SPARQL endpoints. Then, each URI was tested by using the `application/rdf+xml` field in their HTTP header and they all returned a correct RDF document.

- *Availability of the DL.* This criterion assesses the availability of the DL in terms of uptime. It can be measured by using a URI and a monitoring service over a period of time.

$$m_{\text{Avai}} = \frac{\text{Number of successful requests}}{\text{Number of all requests}} \tag{26}$$

The online services were monitored for a period of 7 days with a 5-minute check interval. Only brief interruptions to the service (lasting for a few minutes) were detected.

- *Availability of a public SPARQL endpoint.* This criterion indicates the existence of a publicly available SPARQL endpoint.

$$m_{\text{SPARQL}} = \begin{cases} 1 & \text{SPARQL endpoint publicly available} \\ 0 & \text{otherwise} \end{cases} \tag{27}$$

The BNE and the BnF deploy a Virtuoso[11] server and the BVMC deploys a RDF4J[12] server. No information about the BNB server could be found on its website. The BVMC and the BNB provide a SPARQL editor that assists users to create a query. The BnF, BNB and BVMC provide sample queries as a guide to non-expert users. Some configuration options for the BnF —such as time-out and sponging— requires users to have some expertise. Occasional timeouts were observed when complex queries were submitted to BNB and BnF.

- *Provisioning of an RDF export.* Additionally to the SPARQL endpoint, a RDF data export can be provided to download the whole dataset.

$$m_{\text{Export}} = \begin{cases} 1 & \text{RDF export available} \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

All libraries, with the exception of the BVMC, provide RDF exports as RDF/XML and N-Triples.

- *Support of content negotiation.* This criterion assesses the consistency between the RDF serialization format requested (RDF/XML, N3, Turtle, and N-Triples) and that which is returned.

$$m_{\text{Negot}} = \begin{cases} 1 & \text{Content negotiation supported and correct content types returned} \\ 0.5 & \text{Content negotiation supported but wrong content types returned} \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

All of the libraries failed to deliver at least one of the formats tested, returning HTML by default rather than the format requested: Turtle was not supported by the BVMC and BNB while N-Triples failed in the case of BnF and BNE.

- *Linking HTML sites to RDF serializations.* HTML pages can be linked to RDF serializations by adding a tag to the HTML header with the pattern `<link rel="alternate" type="{content type}" href="{URL}">`.

$$m_{\text{HTMLRDF}} = \begin{cases} 1 & \text{Autodiscovery pattern used at least once} \\ 0 & \text{otherwise} \end{cases} \tag{30}$$

Only the BNB includes such links.

- *Provisioning of repository metadata.* The repository can be described using Vocabulary of Interlinked Datasets (VoID).[52] This criterion indicates whether a machine-readable metadata about the dataset is available.

$$m_{\text{Meta}} = \begin{cases} 1 & \text{Machine-readable metadata available} \\ 0 & \text{otherwise} \end{cases} \tag{31}$$

```
SELECT *
WHERE { ?s owl:versionInfo ?info }
```

Figure 10. SPARQL query retrieving the dataset version.

The BVMC and the BNB report the title, number of triples and vocabularies while the BnF and BNE report the version of the ontology —see Fig. 10.

Discussion. *Accessibility* is also generally high since SPARQL endpoints are provided and they run without significant outages. However, only the BNB and the BVMC provide metadata describing the dataset.

3.12 License

Definition. Licensing is defined as *the granting of permission for a consumer to reuse a dataset under defined conditions*.[43]

Providing a clear and open license is fundamental in order to promote the reuse of a dataset. Licensing can be provided as text in the official website and as machine-readable metadata in the dataset.

Assessment.

There is only one criterion associated with licensing:

- *Provisioning machine-readable licensing information.* A machine-readable license can be specified by means of the relations `dcterms:licence` and `dcterms:rights` included in either the dataset itself or a separate VoID file.

$$m_{\text{macLicense}} = \begin{cases} 1 & \text{machine-readable licensing} \\ & \text{information available} \\ 0 & \text{otherwise} \end{cases} \tag{32}$$

Data are distributed under a Creative Commons CC0 4.0 (Universal Public Domain[13]) with the exception of the BnF repository, whose open license enforces attribution.[53]

Discussion. *Licensing* information is always published, but only the BNB distributes machine-readable licensing information.

3.13 Interlinking

Definition. Interlinking is the extent *to which entities that represent the same concept are linked to each other, be it within or between two or more data sources*.[43]

Interlinking is the key for enriching a dataset: by interlinking a dataset with external repositories, new knowledge can be created. For instance, creating a link to GeoNames provides a well defined and curated knowledge.

Table 7. Number of external `owl:sameAs` links per dataset.

dataset	number	percentage
BNE	522,015	0.06
BnF	13,291,635	0.39
BVMC	63,011	0.04
BNB	4,000,000	0.17

Assessment.

The interlinking dimension measures the number and validity of external links:

- *Interlinking via owl:sameAs*. This score is obtained as the rate of instances having at least one `owl:sameAs` triple pointing to an external resource. Let I_g be the set of instances in g, we can state:

$$m_{\text{Inst}} = \frac{|\{x \epsilon I_g | \exists \{x, \text{sameAs}, y\} \epsilon g \wedge y \epsilon U_g^{ext}\}|}{|I_g|} \tag{33}$$

  The figures are shown in Table 7. We have identified a number of properties that are also used to connect a repository with external sources, such as `umbel:isLike`, `skos:closeMatch`, `skos:exactMatch`, and `dcterms:subject`. The total number of external links is shown in Table 8.
- *Validity of external URIs*. Linking to external resources can lead to invalid links. Given a list of URIs, this criterion checks if there is a timeout or error. Let A be the set of external URIs, then:

$$m_{\text{URIs}} = \frac{|\{x \epsilon A \wedge x \text{ is resolvable}\}|}{|A|} \tag{34}$$

  The number of timeouts and HTTP errors were computed for a random sample of 2,000 URIs defined with the `owl:sameAs` relation and retrieved from their SPARQL endpoints.

Discussion. Although a non-negligible fraction (up to one third) of the instances in every dataset is connected to external repositories, further work is needed in all cases to increase the *interlinking* dimension.

4 Conclusions

Linked open data repositories published by digital libraries have not been assessed by means of a

Table 8. Number of external links to open knowledge bases per repository.

Target	URI	BNE	BnF	BNB	BVMC
BNB	bnb.data.bl.uk	-	-	-	1,626
BNE	datos.bne.es	-	-	-	6,017
BnF	data.bnf.fr	114,114	-	-	5,672
DBpedia	dbpedia.org/resource	52,936	141,244	-	21,749
DDC	dewey.info	-	99,572	-	-
Europeana	www.europeana.eu	-	-	-	46,173
GeoNames	sws.geonames.org	-	-	3,256,918	-
GND	d-nb.info/gnd	157,910	-	-	-
IdRef	www.idref.fr	125,116	1,030,807	-	-
IMSLP	imslp.org		5,546	-	-
ISNI	isni-url.oclc.nl/isni	230,183	1,516,654	1,491,245	5,619
Lexvo	lexvo.org/id/iso639-3	-	-	3,993,674	-
LOC	id.loc.gov/	179,500	342	1,491,245	-
Music brainz	musicbrainz.org	-	42,381	-	-
UK Ref	reference.data.gov.uk	-	-	3,238,656	-
VIAF	viaf.org/viaf	555,097	2,725,515	2,500,000	8,538
Wikidata	www.wikidata.org	-	310.724	-	5,869
Wikipedia	es.wikipedia.org/wiki	48,040	-	-	-
Youtube	www.youtube.com	-	-	-	1,180

quantitative evaluation so far. Based on previous research, we adapted the methodology for LOD repositories to digital libraries. The criteria have been enhanced with a new criterion that checks the number of duplicates.

The application of the methodology described in Section 3 provides a comprehensive picture of the quality achieved by the linked open data repositories created by digital libraries. Four relevant repositories have been evaluated as regards 35 criteria covering 11 dimensions.

The figures in Table 9 are useful to select the dataset that best fits a specific purpose. For instance, if the most relevant aspects for an institution are licensing and interlinking, then the BnF might be the first choice in order to enrich a collection.

Future work to be explored includes the further generalization and automation of the evaluation procedures and the redefinition of some criteria. In addition, possible vocabularies in order to publish the results as LOD will be explored.

A List of prefixes

The prefixes in Table 10 are used to abbreviate namespaces throughout this paper.

Acknowledgements

This work has been partially supported by ECLIPSE-UA RTI2018-094283-B-C32 (Spanish Ministry of Education and Science)

A.1 References

References

1. Berners-Lee T, Hendler J and Lassila O. The semantic web in scientific american. *Scientific American Magazine* 2001; 284.
2. World Wide Web Consortium (W3C). Resource Description Framework (RDF). http://www.w3.org/RDF, 2014. [Online; accessed 10-July-2018].
3. World Wide Web Consortium (W3C). SPARQL Query Language for RDF. https://www.w3.org/TR/rdf-sparql-query/, 2008. [Online; accessed 10-July-2018].
4. Marden J, Li-Madeo C, Whysel N et al. Linked open data for cultural heritage: evolution of an information technology. In Albers MJ and Gossett K (eds.) *Proceedings of the 31st ACM international conference on Design of communication, Greenville, NC, USA, September 30 - October 1, 2013*. ACM, pp. 107–112. DOI:10.1145/2507065.2507103. URL https://doi.org/10.1145/2507065.2507103.
5. Candela G, Escobar P, Carrasco RC et al. A linked open data framework to enhance the discoverability and impact of culture heritage. *Journal of Information Science* 0; 0(0): 0165551518812658. DOI:10.1177/0165551518812658. URL https://doi.org/10.1177/0165551518812658. https://doi.org/10.1177/0165551518812658.
6. Jett J, Cole TW, Han MK et al. Linked open data (LOD) for library special collections. In *2017 ACM/IEEE Joint Conference on Digital Libraries, JCDL 2017, Toronto, ON, Canada, June 19-23, 2017*. pp. 309–310. DOI: 10.1109/JCDL.2017.7991604. URL https://doi.

Table 9. Summary of results.

Dimension	Criterion	BNE	BnF	BNB	BVMC
Accuracy	Syntactic validity of RDF documents	1	1	1	1
	Syntactic validity of literals	1	1	0.9982	1
	Semantic validity of triples	1	1	1	1
	Check of duplicate entities	0.9945	0.9957	0	0.9671
Trustworthiness	On library level	0.25	0.25	0.25	0.25
	On statement level	0	0	0	0
	Using unknown and empty values	0	0	0	0
Consistency	Consistency of schema restrictions during insertion of new statements	0	0	0	0
	Consistency of statements with respect to class constraints	1	1	1	1
	Consistency of statements with respect to relations constraints	0.98	1	1	1
Relevancy	Creating a ranking of statements	0	0	0	0
Completeness	Schema completeness	0.7	0.8	0.65	1
	Column completeness	0.42	0.42	0.34	0.52
	Population completeness	0.59	0.63	0.35	0.14
Timeliness	Frequency	0.5	0.5	0.5	0.5
	Specification of the validity period of statements	0	0	0	0
	Specification of the modification date of statements	0	1	0	0
Ease of understanding	Description of resources	0.93	0.91	0.89	0.92
	Labels in multiple languages	0	1	0	0
	Understandable RDF serialization	1	1	1	1
	Self-describing URIs	1	1	0	1
Interoperability	Avoiding blank nodes and RDF reification	1	1	1	1
	Provisioning of several serialization formats	1	1	1	1
	Using external vocabulary	0.53	0.69	0.90	1
	Interoperability of proprietary vocabulary	0.81	0.85	0.35	1
Accessibility	Dereferencing possibility of resources	1	1	1	1
	Availability of the repository	0.86	0.99	1	0.99
	Availability of a public SPARQL endpoint	1	1	1	1
	Provisioning of an RDF export	1	1	1	0
	Support of content negotiation	0.5	0.5	0.5	0.5
	Linking HTML sites to RDF serializations	0	0	1	0
	Provisioning of metadata	0	0	1	1
Licensing	Provisioning machine-readable licensing information	0	0	1	0
Interlinking	Interlinking via owl:sameAs	0.07	0,39	0,17	0.04
	Validity of external URIs	1	1	1	1

org/10.1109/JCDL.2017.7991604.

7. Mika P, Tudorache T, Bernstein A et al. (eds.). *The Semantic Web - ISWC 2014 - 13th International Semantic Web Conference, Riva del Garda, Italy, October 19-23, 2014. Proceedings, Part I, Lecture Notes in Computer Science*, volume 8796. Springer, 2014. ISBN 978-3-319-11963-2. DOI:10.1007/978-3-319-11964-9. URL https://doi.org/10.1007/978-3-319-11964-9.
8. Auer S, Bizer C, Kobilarov G et al. Dbpedia: A nucleus for a web of open data. In Aberer K, Choi K, Noy NF et al. (eds.) *The Semantic Web, 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007, Busan, Korea, November 11-15, 2007., Lecture Notes in Computer Science*, volume 4825. Springer, pp. 722–735. DOI:10.1007/978-3-540-76298-0_52. URL https://doi.org/10.1007/978-3-540-76298-0_52.
9. Tanon TP, Vrandecic D, Schaffert S et al. From freebase to wikidata: The great migration. In Bourdeau J, Hendler J, Nkambou R et al. (eds.) *Proceedings of the 25th International Conference on World Wide Web, WWW 2016, Montreal, Canada, April 11 - 15, 2016*. ACM, pp. 1419–1428. DOI:10.1145/2872427.2874809. URL https://doi.org/10.1145/2872427.2874809.
10. Rebele T, Suchanek FM, Hoffart J et al. YAGO: A multilingual knowledge base from wikipedia, wordnet, and geonames. In Groth PT, Simperl E, Gray AJG et al. (eds.) *The Semantic Web - ISWC 2016 - 15th International Semantic Web Conference, Kobe, Japan, October 17-21, 2016, Proceedings, Part II, Lecture Notes in Computer Science*, volume 9982. pp. 177–185. DOI:10.1007/978-3-319-46547-0_19. URL https://doi.org/10.1007/978-3-319-46547-0_19.

Table 10. Common prefixes used to designate RDF vocabularies.

prefix	URI
bibo	http://purl.org/ontology/bibo/
blt	http://www.bl.uk/schemas/bibliographic/blterms#
bneonto	http://datos.bne.es/def/
bnfroles	http://data.bnf.fr/vocabulary/roles/
dcmitype	http://purl.org/dc/dcmitype/
dcterms	http://purl.org/dc/terms/
foaf	http://xmlns.com/foaf/0.1/
frbr	http://iflastandards.info/ns/fr/frbr/frbrer/
isbd	http://iflastandards.info/ns/isbd/elements/
owl	http://www.w3.org/2002/07/owl#
prov	http://www.w3.org/ns/prov#
rdac	http://rdaregistry.info/Elements/c/
rdafrbr	http://rdvocab.info/uri/schema/FRBRentitiesRDA
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs	http://www.w3.org/2000/01/rdf-schema#
schema	http://schema.org/
skos	http://www.w3.org/2004/02/skos/core#
umbel	http://umbel.org/umbel/sc/
void	http://www.w3.org/TR/void#
wdt	http://www.wikidata.org/entity/
wd	http://www.wikidata.org/entity/
xsd	http://www.w3.org/2001/XMLSchema#

11. Ehrlinger L and Wöß W. Towards a definition of knowledge graphs. In Martin M, Cuquet M and Folmer E (eds.) *Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems - SEMANTiCS2016 and the 1st International Workshop on Semantic Change & Evolving Semantics (SuCCESS'16) co-located with the 12th International Conference on Semantic Systems (SEMANTiCS 2016), Leipzig, Germany, September 12-15, 2016.*, *CEUR Workshop Proceedings*, volume 1695. CEUR-WS.org, pp. –. URL http://ceur-ws.org/Vol-1695/paper4.pdf.
12. Adamou A, Brown S, Barlow H et al. Crowdsourcing linked data on listening experiences through reuse and enhancement of library data. *Int J on Digital Libraries* 2019; 20(1): 61–79. DOI:10.1007/s00799-018-0235-0. URL https://doi.org/10.1007/s00799-018-0235-0.
13. Achichi M, Lisena P, Todorov K et al. DOREMUS: A graph of linked musical works. In *The Semantic Web - ISWC 2018 - 17th International Semantic Web Conference, Monterey, CA, USA, October 8-12, 2018, Proceedings, Part II*. pp. 3–19. DOI:10.1007/978-3-030-00668-6_1. URL https://doi.org/10.1007/978-3-030-00668-6_1.
14. Debattista J, Lange C, Auer S et al. Evaluating the quality of the LOD cloud: An empirical investigation. *Semantic Web* 2018; 9(6): 859–901. DOI:10.3233/SW-180306. URL https://doi.org/10.3233/SW-180306.
15. Färber M, Bartscherer F, Menne C et al. Linked data quality of dbpedia, freebase, opencyc, wikidata, and YAGO. *Semantic Web* 2018; 9(1): 77–129. DOI:10.3233/SW-170275. URL https://doi.org/10.3233/SW-170275.
16. Joint Steering Committee for Revision of AACR. *Anglo-American Cataloguing Rules, Second Edition*. American Library Association Canadian Library Association, 1998.
17. Standing Committee of the IFLA Cataloguing Section. *International Standard Bibliographic Description (ISBD)*. De Gruyter Saur: IFLA, 2011.
18. IFLA. *IFLA Study Group on the FRBR. Functional Requirements for Bibliographic Records*. München: IFLA Series on Bibliographic Control, 1998.
19. RDA Steering Committee. RDA Toolkit: Resource Description and Access. http://www.rdatoolkit.org, 2012. [Online; accessed 19-November-2018].
20. Aalberg T and Zumer M. Looking for entities in bibliographic records. In Buchanan G, Masoodian M and Cunningham SJ (eds.) *Digital Libraries: Universal and Ubiquitous Access to Information, 11th International Conference on Asian Digital Libraries, ICADL 2008, Bali, Indonesia, December 2-5, 2008. Proceedings*, *Lecture Notes in Computer Science*, volume 5362. Springer, pp. 327–330. DOI:10.1007/978-3-540-89533-6_36. URL https://doi.org/10.1007/978-3-540-89533-6_36.
21. Vila-Suero D, Villazón-Terrazas B and Gómez-Pérez A. datos.bne.es: A library linked dataset. *Semantic Web*

2013; 4(3): 307–313. DOI:10.3233/SW-120094. URL https://doi.org/10.3233/SW-120094.

22. Candela G, Escobar P, Carrasco RC et al. Migration of a library catalogue into RDA linked open data. *Semantic Web* 2018; 9(4): 481–491. DOI:10.3233/SW-170274. URL https://doi.org/10.3233/SW-170274.
23. Waagmeester A, Willighagen EL, Queralt-Rosinach N et al. Linking wikidata to the rest of the semantic web. In *Proceedings of the 9th International Conference Semantic Web Applications and Tools for Life Sciences, Amsterdam, The Netherlands, December 5-8, 2016*. URL http://ceur-ws.org/Vol-1795/paper46.pdf.
24. Wikidata. SPARQL federation input/Archive. https://www.wikidata.org/wiki/Wikidata:SPARQL_federation_input/Archive, 2017. [Online; accessed 10-July-2018].
25. Sim SE, Easterbrook SM and Holt RC. Using benchmarking to advance research: A challenge to software engineering. In *Proceedings of the 25th International Conference on Software Engineering, May 3-10, 2003, Portland, Oregon, USA*. pp. 74–83. DOI:10.1109/ICSE.2003.1201189. URL https://doi.org/10.1109/ICSE.2003.1201189.
26. Heckman SS and Williams L. On establishing a benchmark for evaluating static analysis alert prioritization and classification techniques. In *Proceedings of the Second International Symposium on Empirical Software Engineering and Measurement, ESEM 2008, October 9-10, 2008, Kaiserslautern, Germany*. pp. 41–50. DOI:10.1145/1414004.1414013. URL https://doi.org/10.1145/1414004.1414013.
27. Spahiu B, Maurino A and Meusel R. Topic profiling benchmarks in the linked open data cloud: Issues and lessons learned. *Semantic Web* 2019; 10(2): 329–348. DOI:10.3233/SW-180323. URL https://doi.org/10.3233/SW-180323.
28. Piscopo A. Wikidata:Requests for comment/Data quality framework for Wikidata. https://www.wikidata.org/wiki/Wikidata:Requests_for_comment/Data_quality_framework_for_Wikidata, 2016. [Online; accessed 11-February-2018].
29. Radulovic F, Mihindukulasooriya N, García-Castro R et al. A comprehensive quality model for linked data. *Semantic Web* 2018; 9(1): 3–24. DOI:10.3233/SW-170267. URL https://doi.org/10.3233/SW-170267.
30. Carrasco MH, Luján-Mora S, Maté A et al. Current state of linked data in digital libraries. *J Information Science* 2016; 42(2): 117–127. DOI:10.1177/0165551515594729. URL https://doi.org/10.1177/0165551515594729.
31. Mitchell ET. Library linked data: Early activity and development. *Library Technology Reports* 2016; 52(1): 5–13. DOI:10.5860/ltr.52n1. URL http://dx.doi.org/10.5860/ltr.52n1.
32. Shen G and Liu G. The selection of benchmarking partners for value management: An analytic approach. *International Journal of Construction Management* 2014; 7. DOI:10.1080/15623599.2007.10773099.
33. IFLA Information Technology Section ; IFLA Semantic Web Special Interest Group ; Bibliothèque nationale de France. *We grew up together: data.bnf.fr from the BnF and Logilab perspectives*. Paris, Bibliothèque nationale de France, Petit auditorium: IFLA Information Technology Section ; IFLA Semantic Web Special Interest Group ; Bibliothèque nationale de France, 2014. URL http://ifla2014-satdata.bnf.fr/program.html.
34. Diane Hillmann, Gordon Dunsire, Jon Phipps. FRBR Entities for RDA vocabulary. http://rdvocab.info/uri/schema/FRBRentitiesRDA, 2014.
35. Bibliothèque nationale de France. Subject reference systems. RAMEAU. http://www.bnf.fr/en/professionals/anx_cataloging_indexing/a.subject_reference_systems.html, 1980.
36. British Library. Basic RDF/XML. "http://www.bl.uk/bibliographic/datafree.html#basicrdfxml", 2014. [Online; accessed 8-November-2018].
37. RDA Steering Committee. RDA Registry. http://www.rdaregistry.info/, 2015. [Online; accessed 11-February-2018].
38. Wang RY and Strong DM. Beyond accuracy: What data quality means to data consumers. *J of Management Information Systems* 1996; 12(4): 5–33. URL http://www.jmis-web.org/articles/1002.
39. Beall J. Metadata and data quality problems in the digital library. *J Digit Inf* 2005; 6(3). URL http://journals.tdl.org/jodi/article/view/65.
40. World Wide Web Consortium (W3C). W3C RDF Validation Service. https://www.w3.org/RDF/Validator/, 2006.
41. Gordon Dunsire. ISBD elements. http://metadataregistry.org/schemaprop/show/id/2128.html, 2015. [Online; accessed 4-April-2018].
42. Online Computer Library Center. The Virtual International Authority File. https://viaf.org/, 2012.
43. Zaveri A, Rula A, Maurino A et al. Quality assessment for linked data: A survey. *Semantic Web* 2016; 7(1): 63–93. DOI:10.3233/SW-150175. URL https://doi.org/10.3233/SW-150175.
44. World Wide Web Consortium (W3C). PROV-O: The PROV Ontology. https://www.w3.org/TR/prov-o/, 2013. [Online; accessed 1-August-2018].

45. Mecella M, Scannapieco M, Virgillito A et al. Managing data quality in cooperative information systems. In *On the Move to Meaningful Internet Systems, 2002 - DOA/CoopIS/ODBASE 2002 Confederated International Conferences DOA, CoopIS and ODBASE 2002 Irvine, California, USA, October 30 - November 1, 2002, Proceedings*. pp. 486–502. DOI:10.1007/3-540-36124-3_28. URL https://doi.org/10.1007/3-540-36124-3_28.

46. Shreeves SL, Knutson E, Stvilia B et al. Is quality metadata shareable metadata? the implications of local metadata practices for federated collections.

47. Cooper MD and Chen H. Predicting the relevance of a library catalog search. *JASIST* 2001; 52(10): 813–827. DOI:10.1002/asi.1140. URL https://doi.org/10.1002/asi.1140.

48. Pipino L, Lee YW and Wang RY. Data quality assessment. *Commun ACM* 2002; 45(4): 211–218. DOI: 10.1145/505248.5060010. URL http://doi.acm.org/10.1145/505248.5060010.

49. Gonçalves MA, Moreira BL, Fox EA et al. ”what is a good digital library?” - A quality model for digital libraries. *Inf Process Manage* 2007; 43(5): 1416–1437. DOI:10.1016/j.ipm.2006.11.010. URL https://doi.org/10.1016/j.ipm.2006.11.010.

50. World Wide Web Consortium (W3C). Notation3 (n3): A readable rdf syntax. ”https://www.w3.org/TeamSubmission/n3/”, 2011. [Online; accessed 13-November-2018].

51. World Wide Web Consortium (W3C). Semantic Integration & Interoperability Using RDF and OWL. https://www.w3.org/2001/sw/BestPractices/OEP/SemInt/, 2005. [Online; accessed 04-September-2019].

52. World Wide Web Consortium (W3C). Describing linked datasets with the void vocabulary. https://www.w3.org/TR/void/, 2011. [Online; accessed 19-February-2018].

53. Etalab. Open platform for french public data. http://data.bnf.fr/docs/Licence-Ouverte-Open-Licence-ENG.pdf, 2011. [Online; accessed 1-March-2018].