

Towards a Methodology for Publishing Library Linked Data

Dydimus Zengenene¹, Vittore Casarosa², and Carlo Meghini²

¹ DILL International Master, University of Parma, Italy

² Istituto di Scienza e Tecnologie della Informazione del CNR, Pisa, Italy

Abstract. It is argued that linked data are becoming increasingly necessary for libraries and related institutions, such as galleries, museums and archives. Though libraries are potentially crucial players in the linked data movement, very often there is lack of knowledge among librarians on how to publish linked data. This paper presents the results of a master thesis whose main aim was to empower libraries to take an effective part in publishing linked data thereby contributing to building the semantic web, in order to improve the general services which they offer. In a narrower sense, this research aims to draw a methodology applicable to the library domain in publishing linked data. A 15-step methodology is presented and illustrated in some detail.

Keywords: linked data, library data.

1 Introduction

Due to their traditional role as the curators of valuable information, and their expertise in metadata generation and management, libraries or, more in general, the so called memory institutions (libraries, archives and museums) are in a unique position of providing trusted metadata for resources of cultural value. Libraries and related institutions are therefore expected or forced by circumstances to be key players in building the new generation of the web called the semantic web or the web of data. [2] notes that Libraries have already taken a leading role in the application of semantic web technologies because they own well described collections of objects.

Because of its ubiquity, scalability and simplicity, the web is recognized as the ideal medium to transform the way data is discovered, accessed, integrated and used. In that regard linked data can be defined as a set of principles and technologies that harness the ethos and infrastructure of the web to enable data sharing and reuse [8]. These concepts are penetrating also in the domain of Library and Information Science (LIS) and related disciplines, and [4] boldly states that “...of all information communities, libraries are in the best position to transition their data into Linked Data because the basic elements already exist in their catalogues”.

The value of Linked Data for libraries derives mainly from the navigation possibilities offered by Linked Data, which provide a seamless information space that can be explored by following “typed” links (i.e. links with a known meaning) much in the

same way as the “normal” web can be explored by following “anonymous” links. Links between libraries and non-library services such as Wikipedia, GeoNames, MusicBrainz, the BBC, and The New York Times will connect local collections into the larger universe of information on the Web. In addition, the use of globally unique identifiers (URIs) to designate works, places, people, events, subjects, and other objects or concepts of interest, allows libraries to increase their presence on the Web, by having their resources cited across a broad range of data sources. The use of unique identifiers also allows the description of a resource to be tailored to specific communities such as museums, archives, galleries, and audiovisual archives. Linked Data are represented in a structured way (RDF), which increases the visibility of library content towards crawling and relevancy algorithms of search engines and social networks.

Though libraries can potentially be crucial players in the linked data movement, very often there is lack of knowledge among librarians on how to publish linked data. The World Wide Web Consortium (W3C) has established a “Library Linked Data Incubator Group”, whose main aim was “to help increase global interoperability of library data on the Web, by bringing together people involved in Semantic Web...”. The final report of the group [12] points out that there are several reasons that make it a great challenge for libraries to successfully participate in the movement. One of the cited reasons is the complexity and variety of available vocabularies in libraries, their overlapping coverage, derivative relationships and alignments. The problem is exacerbated by the fact that most library and information professionals are unfamiliar with linked data sets and vocabularies that can be of use in the library domain because these data sets have been developed in the semantic web research community. There are efforts to participate in the movement, however most such library projects define themselves as prototypes.

The broad aim of this research is to empower libraries to take an effective part in publishing linked data thereby contributing to building the semantic web, in order to improve the general services which they offer. In a narrower sense, this research aims to draw a methodology applicable to the library domain in publishing linked data. In the remainder of this paper we provide a brief description of the basic principles underlying Linked Data (Section 2), then we describe briefly the methodology used to identify a 15-step process for publishing library data as linked data (Section 3), then we describe in some detail each of those steps (Section 4) and finally we provide some concluding remarks about the application of the process (Section 5).

2 Basics on RDF and Linked Data

Linked data is associated with a new generation of the web called the semantic web, the web of data or sometimes web 3.0. [8] define Linked Data as a set of “...best practices for publishing and connecting structured data on the Web”. The World Wide Web Consortium (W3C) shortly defines the semantic web as the web of data, which is capable of providing a common framework that allows data to be shared and reused across application, enterprise, and community boundaries.

Linked Data is not about creating a different Web, but rather about enhancing the Web through the addition of structured information. This structured information, expressed using technologies such as URIs for identifiers of resources and RDF [14] for the description of their properties, specifies relationships between things, that can then be used for navigating between, or integrating, information from multiple sources.

RDF is based on the idea of identifying things (resources and their properties) using Web identifiers (called Uniform Resource Identifiers, or URIs), and describing resources in terms of simple statements asserting properties of a resource (sometimes called attributes) by providing a property and its and its associated values. This enables RDF to describe resources as a graph, where a node representing a resource (called the subject) can have a number of arcs, each one associated with a specific property, leading to another node (called the object) representing the value of that property. The value of a property (the object node) can be either a terminal string, in which case the graph ends there, or the URI of another resource, in which case the graph can continue with the properties of this new resource, which becomes the subject of a new set of statements.

In addition to providing a way to express simple statements about resources, using named properties and values, RDF also provides, through the RDF Schema, the ability for user communities to define the vocabularies (terms) that they intend to use in those statements: Those vocabularies (sometimes, maybe improperly, called ontologies) indicate the specific kinds or classes of resources that are going to be described, and the specific properties that are going to be used in describing those resources. RDF Schema provides the facilities needed to describe such classes and properties, and to indicate which classes and properties are expected to be used together (for example, to say that the property “jobTitle” will be used in describing a resource belonging to the class “Person”).

In essence, Linked Data is a general way to organize information as “clouds of data” (RDF graphs) describing resources in specific application domains, which can be connected together in order to enrich each other and provide additional information. The resulting “global cloud” can be navigated to get the description of the resources and, following their properties and connections, to discover unexpected aspects or relationships of the objects of interest.

In the Linked Data world, it is important to distinguish between a resource, i.e. a real world object or concept, which not necessarily needs to be digital (and therefore downloadable from the Internet) and the description of that resource, which usually is digital and can be retrieved by navigating the Linked Data. It is common to provide the description of the resource in two ways, one expressed in HTML for use by a human through a browser, and one expressed in RDF/XML for use by an application e program. In this way, a resource usually leads to the definition of at least three URIs, the one that represents the real world object, the one that represents its HTML representation and the one that represents its RDF/XML representation.

3 Methodology

The data gathered by the W3C Incubator Group of Library Linked Data [12] was a starting point for this work. Even though this data was not collected to capture the

process, it was analysed identifying specific process issues that linked data publishers came across. These issues were used in conjunction with project reports to come up with questions which were used as a guideline for in-depth interviews. This approach was considered in order to avoid redoing the work that the W3C had already done but instead built on data which was collected by the W3C when it was preparing its report on Libraries and linked data. While W3C concentrated on the state of affairs in libraries, this work was concerned with implementation issues. This gave a different interpretation of the same data. However the data which was collected by the W3C was insufficient to meet the needs of this work since it was collected with a different intention. To cover for the shortfalls, we analyzed project reports of selected cases and prepared questions which were used for interviews. Books, opinion papers on linked data and web based resources were used to inform some sections of the work which both the data and the interviews could not fill. The complete results of the research are given in [13]. In this paper, we summarize the methodology, discussing its steps in some detail.

4 The Recipe

This section summarizes a 15 step practical methodology which can be followed for publishing linked data. As in libraries there are various datasets which can potentially be published as linked data, it should be clear that the exact workflow may vary depending on the nature of the data to be published, and also the fact that linked data is an evolving technology makes it difficult to define the “ideal workflow”. However we can safely say that the library services and systems will become part of the “new web” if the library will manage to publish the following data: knowledge organization systems (classification schemes, thesauri), authority files, digital contents and their descriptions, catalogue data including circulation data sets. All these datasets should have links within themselves and should establish outgoing links to many other web resources, in order to attract many incoming links in what [3] calls “Web Centric Cataloguing”.

4.1 Step 1: Motivation

It is important to know what is motivating a library to adopt linked data technologies. This implies understanding what linked data technology can do to improve your legacy system, how much it can improve users search and browsing experiences and how much it empowers them to contribute in enriching the information collection and in collaborating among themselves. The best motivations seem to come from comparing the current system against the potential of linked data enabled systems. Publishing the data on the web will make the organization more visible, and we may assume that this is what helps the library to better achieve its mission. This is supported by the fact that most libraries involved with large linked data projects are National Libraries and academic libraries, whose services may benefit by reaching a much wider public. All these considerations have to be written down clearly and convincingly in non technical jargon in order to use the document in the next step.

4.2 Step 2: Management Approval

Using the motivations in step one it is important to gain stakeholder approval and support. Management support to the project is important, given that so far no financial gains have been reported. On the contrary, some libraries have even abandoned their policy of selling metadata in favor of making them available for free as linked data. This is a management decision which one interviewee defined as “a great step”. The benefits are also tangible since they can be measured by key library indicators, like user satisfaction, system improvement (precision, recall), innovative use and collaboration. This however depends on the ability for librarians to clearly view these benefits as a direct improvement of their services, clearly articulate the benefits and sell the idea to the stakeholders giving assurance on such sensitive issues like personal data security and controlled access to licensed resources.

4.3 Step 3: Sorting Out the Legal and Financial Issues

This stage has to do with assessing the rights that the institution has over the data sets. When there is a need to make contractual agreements with the data owners, there is also the need to know if the institution has the legal capacity to enter into such agreements. It is at this stage that license issues can be discussed, deciding what licenses and waivers the institution will grant to the intended users of the datasets. It is possible to have different licenses and waivers on different datasets of the same institution. The question of data licensing is still an open research issue [11]. Following linked data publishing principles does not necessarily mean opening up the data. There is an option for closed linked data. Miller and colleagues argue that the use of Creative Commons Licenses is mostly due to the lack of better alternatives, as pointed out at the site “Bibliographic Wilderness” (2008), which states that catalogue data is not “copyrightable”.

4.4 Step 4: Assessment of Skills and Data Available

At this stage it is convenient to start planning the conversion process. In order to have a precise understanding of what is needed, this stage must be based on a situation analysis, which should include the assessment of the existing skills and the identification of the datasets to be published and the formats in which they are. The options available have to be weighed against the skill that are needed. The fact that several cases reported to have relied on internal resources might suggest that it is not very difficult for motivated people to learn the needed skills “on the job”. However, as noted by Coyle [4] librarians must transform their skills into understanding and managing ontologies, understanding information system design, etc., so that they can communicate with technical experts in carrying out such projects. Required skills:

- Information systems skills: these are of value in downloading, installing and configuring systems, databases (especially triple stores), and other servers, writing and reading XML and RDF.

- Metadata skills: knowledge of the cataloguing process and the meaning of the metadata elements.
- Modeling: understanding the data structure and coming up with the correct algorithms of converting data from a given structure to RDF.

If the library can hire external experts, the librarians need to communicate with the technical people in order to express their requirements. Therefore they need to:

- Be able to communicate their technical needs by speaking requirement modelling languages like the Universal Modeling Language.
- Be able to describe their ontologies using such vocabularies as FRBR, FRBRoo, CIDOC CRM, FOAF etc.
- Read and understand XML RDF records in various serialization formats to able to evaluate the end result.

Analyzing the data and the systems that hold them gives an overall idea of how complex the task might turn out to be. The diagram in Figure 1 gives an overall view of the conversion process, depending on the type of existing data. Depending on the type of data and the system where they are stored different conversion tools (such as entity extractors, RDF-izers, wrappers, RDB-to-RDF, etc. as indicated in the figure) should be used.

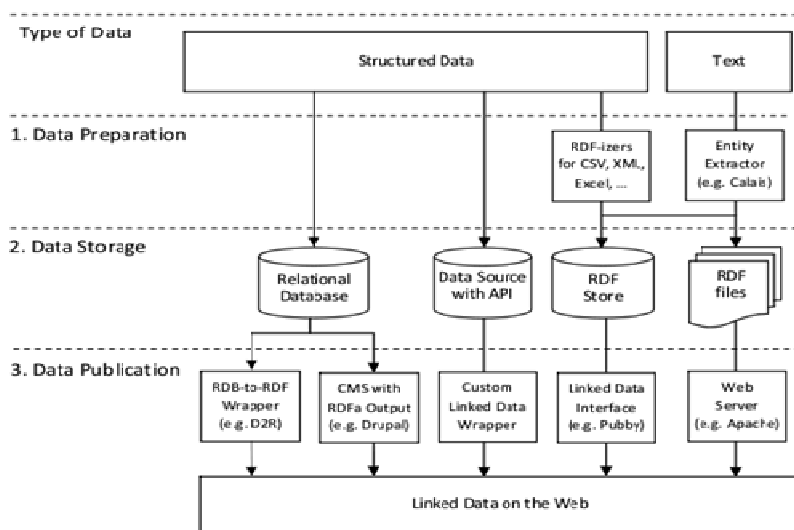


Fig. 1. Publishing workflow (from [8], p. 70)

4.5 Step 5: Tools Assessment and Evaluation

Referring again to Fig. 1, if the data exists in text format there is the need to use entity extractors, which provide RDF annotations on text documents. The resultant RDF

files can be published through a web server or through an RDF triple store. When the data is in other non queryable structured formats (e.g. spreadsheets), conversion takes place using other special tools referred to in the diagram as “RDF-izers” and then served via an RDF triple store. Many bibliographic systems store bibliographic data in queryable relational data format and digital libraries store their content in content management systems (CMS). In both cases there is little need to prepare the data, and for bibliographic data a custom wrapper should be used, while in the second case a CMS with RDFa output should be used. Knowing the data and the possible flow will help to know the skills, the experience and the tools that are needed. For example, using a custom wrapper is considered to be the most complex situation which cannot be accomplished without some programming skills [8].

4.6 Step 6: Dataset Analysis

This is a critical stage where most problems are likely to be faced. In addition to knowing the system that hosts the data and the data format, it is important to know also the level of description and the metadata elements sets in use. In the case of descriptions, very often they might be incomplete, or they might be based on systems that do not support linked data, or they might not follow any particular metadata standard. In the case of projects that involve aggregating data originating from different systems (e.g. from different institutions), or different systems like catalogues and digital library systems, there is the likelihood that these systems will be using different metadata standards (e.g. MARC and Dublin Core) and in this case there will be the need to map one set to the other.

4.7 Step 7: URI Assignment

Each resource in the dataset has to be identified by a unique URI, created according to the following guidelines:

- Use HTTP URIs so that they are dereferenceable.
- Ensure that the URIs are from a namespace that you control.
- Make sure your URIs do not carry implementation details which can change over time.
- It is advisable to use meaningful natural keys in URIs as unique identifiers of resources; for example, books can be identified by using the ISBN number instead of primary keys in the local database.

As described before, one resource in a dataset usually leads to the creation of at least three URIs, the one that represents the real world object, the one that represents its HTML representation and the one that represents its RDF/XML representation. In the world of Linked Data there are three common ways to distinguish these URIs. The most common one, used by dbpedia, Europeana and others, is the use of the terms “resource”, “page” and “data” in their URIs, to indicate the real world object, its HTML representation and its RDF/XML representation respectively [8]. Another alternative is to use the terms “id”, “pages” and “data” on their URIs. Finally, the third

alternative is to use the file extensions, using no extension for the resource, “html” for the HTML representation and “rdf” for the RDF representation.

Whereas making URIs for resources that you control involves making a choice of the representation that you prefer, external URIs do not give such freedom. In many cases, different institutions refer to the same resource using different URIs. If you have the same resource in your data set, which URI should you use? One alternative is to define your own, and then declare its equivalency with the others by using the “owl:sameAs” property. Another alternative is to use an existing one, taking into consideration the authority and reliability of the institution that defined it. For domain specific datasets one may find domain specific authorities like GeoNames for geographic names, VIAF for author names in libraries, FOAF for people etc. For general datasets more general authorities like dbpedia may be more useful.

4.8 Step 8: Vocabulary Modeling

Vocabulary modelling has to do with creating controlled terminology giving explicit meaning to the concepts in your dataset, and this is a key process in linked data. The emphasis is on the use of already existing vocabularies evaluated using the four criteria as noted by [8] (p.62). A vocabulary of choice should:

- be widely used to ensure widespread use of your dataset,
- be actively maintained according to a clear governance process,
- cover enough of your dataset to justify its terms, and
- be expressive enough to suit your particular requirements.

There is no single place to find vocabularies. From the literature, libraries are recommended to consider their own domain models as a starting point. These include FRBR, FRAD, and the more recent FRBRoo [4]. Alternatively, it is recommended to find suitable vocabularies by asking what other people in the same domain are using. Swoogle, Sindices and Taxonomy warehouse are some of the sources to find vocabularies [1]. Also the W3C supplementary report on data sets, vocabularies and metadata elements sets is a good starting point to see what others have used before.

A good knowledge of your existing data will definitely help in the choice of a vocabulary. Does your data include People, Places, Books, Journals, Films, Authors, Musicians, Concepts, Photos, Comments, Reviews and so on. This knowledge makes it possible to have quick choices of vocabularies which are widely used at the moment. For example, Geonames would be among the first to consider for Places, and FOAF would be the one to consider for People [7]. In the most common cases some new terms are introduced as different existing vocabularies are brought together to satisfy one’s need. Such new terms are called proprietary terms. They should be made deferenceable to ensure that applications and other users can retrieve their definitions in order to know their meaning and perhaps reuse them. In the worst case one might be forced to build a new vocabulary from scratch. One major reason why it is always preferable to reuse existing vocabularies is the fact that whenever you create a new vocabulary in the linked data world you have created a data island and have decreased the level of understanding in that domain. Unless you are a very well known authority in that domain, it is likely that your vocabulary will not be used by someone else.

Creating a vocabulary is a very complex process which needs linguistic skills and domain expertise. Creating your own vocabulary should be the very last option in vocabulary modelling. In the event that you create your own vocabularies, consider relating your new vocabulary to known vocabularies making them sub concepts of the known vocabularies.

4.9 Step 9: Generation of RDF Data

If the preceding steps were done well, the process of generating the RDF data should not be a difficult one, as this is the stage of implementation of decisions taken before. At the moment there is a number of tools being developed to help in this process, so that in the future there might be less need to have strong coding skills but rather skills for configuring and customizing would be needed. In many cases, the amount of data is too large to convert existing literal values into URIs manually, and the use of tools to do the process automatically or semi-automatically is mandatory. It is not part of this work to compare and contrast existing tools (some of them are being developed in house by major projects) as each of them has its own strengths and weaknesses. Reviews of these systems can be found easily on the web.

4.10 Step 10: Enriching the Data

This process involves defining news triples that connect resources within the published dataset or to other resources in other external datasets, creating triples that define relationships internally within the dataset (internal links) or relationships with outside resources (outgoing links). For internal links it must be ensured that every part of the dataset is reachable by a crawler when it is following links and therefore each file has to be connected to related files in the same dataset. For outgoing links, it is advisable to start by linking to such datasets as dbpedia, Geonames, Europeana, VIAF and others, which are already well established and stable in the linked data world. This ensures that your dataset is easily discoverable since these are widely linked to by many other datasets. As already stated, points of caution are: (1) URIs must not carry implementation detail which can change over time (e.g., port numbers, server names or php extensions). (2) Use natural keys in URIs as unique identifies of resources (e.g., ISBN number instead of using the primary key in your local database). (3) natural key should be meaningful within the domain (e.g. ISBN is meaningful in the library domain). For external links ensure that you are connecting to datasets that adds value to your set and are reliable. The questions to be asked when linking to external datasets are:

- What is the value of your data in the target dataset?
- To what extent does this add value to the new dataset?
- Is the target dataset and its namespace under stable ownership and maintenance?
- Are the URIs in the dataset stable and not likely to change?
- Are there outgoing links to other datasets?

There is research going on about ranking datasets. Such works include “DING”, a novel two-layer ranking model for the Web of Data [5].

4.11 Step 11: Describing the Data Set

Before publishing, there is the need to provide a description of the dataset, which includes provenance and licensing metadata. In provenance metadata one might describe the history of that dataset, how it was generated and the technical processes that have been undergone to establish the dataset. In license and waiver metadata, one describes how that dataset maybe used by third parties. A good description of the dataset is therefore essential to establish trust among third parties and also to ensure that people will use the dataset according to the conditions set by the publisher.

In addition to making instance data self-descriptive, it is also desirable that data publishers provide metadata describing characteristic of complete data sets, for instance, the topic of a data set and, if possible, more detailed statistics about the data. This information might include label, URI of the dataset, location of SPARQL endpoint, data dumps, last-modified date of the dataset, and change frequency [8] (p. 48). These data may be described using the Vocabulary of Interlinked datasets (void), which is an RDF vocabulary, as in the “datahub” web site [9].

4.12 Step 12: Evaluating the Dataset

Before publishing the data for access by third parties, it is advisable to evaluate it to see how good it is and if it conforms to standards. A possible checklist proposed by [8] (p. 83) is the following.

- Does your data set links to other data sets?
- Do you provide provenance metadata?
- Do you provide licensing metadata?
- Do you use terms from widely deployed vocabularies?
- Are the URIs of proprietary vocabulary terms dereferenceable?
- Do you map proprietary vocabulary terms to other vocabularies?
- Do you provide data set-level metadata?
- Do you refer to additional access methods?

At this stage is also advisable to check if the triples are well expressed. This can be done by using such tools as Vapour Linked Data Validator, RDF: Alerts and Sindice Inspector. In addition, there is another level of evaluating a dataset, which involves evaluating the data quality. This has to do with how you can find the goodness of a dataset and weather it is worth linking to. Some key criteria have been mentioned in the step of selecting a vocabulary to link to. Finally, it is also advisable to use the 5 stars rating system recommended by W3C.

4.13 Step 13: Publishing

The decision of how the data will be published should have been done in Step two, where the workflow was planned, and where the need to provide users with several access points to ensure that the dataset is widely used should have been considered. The most obvious way to publish Linked Data on the Web is to make the URIs that identify data items dereferenceable into RDF descriptions. In addition, various Linked

Open Data providers, including libraries, provide two alternative means of accessing the data, namely via SPARQL endpoints or by providing RDF dumps of the complete data set. In general the system should provide access to both the RDF and HTML representations of the data. This is usually done by configuring 303 redirects in triple stores in response to a client request to access either the HTML representation of an object or its RDF representation.

4.14 Step 14: Incoming Links

Incoming links originate from other datasets, linking into your dataset. Third parties need to be convinced that your dataset is valuable to them so that they can link to it. However it is usually difficult for them to know your value unless you do some sort of marketing and promotion actions. As a starting point a publisher can create triples that link to their own dataset and ask third parties like dbpedia to add those triples to their own dataset [8] (p.64) To continue attracting new links, there might be the need to employ marketing techniques so that the dataset is known by new users and be linked to.

4.15 Step 15: Feedback

Once the data has been published and incoming links solicited, it is important to wait for feedback from the user community so that one can incorporate their views and refine the dataset. From the interviews conducted it was found out that this is a continuous process, even if it is usually very difficult, if not impossible, to know what people are doing with the RDF data. Download counts could give a rough idea of your data usage, and Google Analytics could provide additional information, such as the change in the number of times a certain URI is accessed, the IP addresses which are accessing your data and their geographic location, so that one could derive a rough idea of how the data is being used.

5 Conclusions

Changes brought about by the internet and the web have affected all walks of life and changed business models under which many institutions have been operating. Libraries and the library profession have not been spared. In order to continue being relevant and effective, libraries have to adapt to the latest developments in technology and the resultant user behaviour. Adopting linked data technologies is one such initiative to improve the library presence where today's information is sought (i.e. the web), to improve the functionality of their systems and to promote innovative use of their data. Even though the 15 step approach helps in following the progress and managing the project, it is recommended taking first a small dataset and carrying out the necessary trials and errors until the data can be successfully published, in order to develop the needed skills. Clearly, for large datasets and large projects that need to be monitored and evaluated, planning and accountability tracing is necessary thereby calling for a methodological approach. This procedural approach should hopefully minimize possible mistakes.

References

1. Where to find vocabularies,
http://vocamp.org/wiki/Where_to_find_vocabularies
2. Isaac, A., Kramer, D., van der Meij, L., Wang, S., Schlobach, S., Stapel, J.: Vocabulary matching for book indexing suggestion in linked libraries – A prototype implementation and evaluation. In: Bernstein, A., Karger, D.R., Heath, T., Feigenbaum, L., Maynard, D., Motta, E., Thirunarayan, K. (eds.) ISWC 2009. LNCS, vol. 5823, pp. 843–859. Springer, Heidelberg (2009)
3. Chudnov, D.: What linked data is missing. *Computers in Libraries* 31(8) (2011)
4. Coyle, K.: Making connections. *Library Journal* 134(7), 44–47 (2009)
5. Delbru, R., Toupikov, N., Catasta, M., Tummarello, G., Decker, S.: Hierarchical link analysis for ranking web data. In: Aroyo, L., Antoniou, G., Hyvönen, E., ten Teije, A., Stuckenschmidt, H., Cabral, L., Tudorache, T. (eds.) ESWC 2010, Part II. LNCS, vol. 6089, pp. 225–239. Springer, Heidelberg (2010)
6. Europeana Foundation. The problem of the yellow milkmaid: A business model perspective on open metadata (2011), <http://pro.europeana.eu/documents/858566/2cbf1f78-e036-4088-af25-94684ff90dc5/>
7. Heath, T.: An introduction to linked data (2009),
<http://tomheath.com/slides/2009-02-austin-linkeddata-tutorial.pdf>
8. Heath, T., Bizer, C.: Linked data. Evolving the web into a global data space. *Synthesis Lectures on the Semantic Web*. Morgan & Claypool (2011)
9. Taxonconcept knowledge base, <http://datahub.io/>
10. Jackson, J.: Open source vs. proprietary software. *PCWorld* (2011)
11. Heath, R., Miller, T., Styles, P.: Open data commons, a license for open data. In: *Proc. of LDOW 2008-Linked Data on the Web Workshops*, Beijing, China (2008)
12. Coyle, K., Dunsire, G., Isaac, A., Murray, P., Panzer, M., Schneider, J., Singer, R., Summers, E., Waites, W., Young, J., Zeng, M., Baker, T., Bermes, E.: Library linked data incubator group final report. W3C (October 2011),
<http://www.w3.org/2005/Incubator/11d/XGR-11d-20111025/>
13. Zengenene, D.: A methodology for publishing linked data in the library domain. Master's thesis, International Master in Digital Library Learning (2012)
14. Manola, F., Miller, E.: RDF Primer. Technical report, W3C (February 2004),
<http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>