

Bibliographic Data Towards the Semantic Web: A Review of Key Issues and Recent Experiences

Bibliyografik Verilerin İşlenmesinde Yenilikçi Yaklaşımlar: Semantik Web'e Doğru

Iryna SOLODOVNIK*

Abstract

This article intends to review the underlying concepts and technologies of the Semantic Web and the potential they provide for metadata management covering bibliographic resources. To get closer to a semantic web data space, different libraries are adhering to the initiatives making their traditional Knowledge Organization Systems (KOS) operational on the web through SKOS techniques, as well as releasing bibliographic data under open licenses (open bibliographic data) and publishing it with Linked Data (LD) mechanisms. LD meaningful semantic connections create the Web of Data, a global database representing the first practical step to the Semantic Web. Here interoperable data can be processed independently of application, platform or domain, providing rich retrieval results produced by powerful query languages. From a library perspective, a problem statement is a global promotion within the Library community of understanding and of adoption of Linked Open Data (LOD), of LODe-BD recommendations, as well as releasing bibliographic data as Linked Library Data (LLD). In this way, different bibliographic datasets could become full members of the Semantic Web making interoperable different knowledge datasets of heterogeneous web communities.

Keywords: Semantic web, Simple knowledge organization system (SKOS), Linked (Open) data, Bibliographic data, LOD-e BD, Linked library data

Öz

Son yıllarda, verilerini Web üzerinden paylaşan global Web topluluklarında, bibliyografik, bilimsel ve yönetsel verilerin yönetilmesi bakımından dijital bilgi sistemleri, yeni bilginin üretiminde merkezi bir rol oynamaya başlamıştır. Dijital materyallerin hacmi büyüdükçe ve kullanımları yaygınlaştıkça, dijital geçmiş ve geleceğimizin irtibatlandırılması önemli bir başarı mücadelesi gerektirmektedir. Geleneksel Bilgiyi Organize Edici Sistemler, dijital çevrede veri yayınlanmasının yeni ortaya çıkan yöntemleri, verileri serbest erişime açma ve ilişkilendirerek "açık veri" haline getirme, farklı toplumların yenilikçi yaklaşımlarını destekleyen işbirliği çalışmaları oluşturma, bibliyografik veri tabanı yönetimine özgü yeni yaklaşımlardır. Aynı şekilde, dijital içeriğin düzenlenmesinde gittikçe karmaşıklaşan ortamlara ihtiyaç duyulması, metaveri modellemesinde yeni çığırar açmaktadır. Bu çalışmanın temel amacı, dijital enformasyon ve veri yönetimindeki yeni kuramsal ve pratik kullanımı olan paradigmalardan okuru haberdar etmek ve özellikle de Semantik Web kapsamındaki bibliyografik verilere odaklanmaktır.

Anahtar sözcükler: Semantik web, Metadata, Bibliyografik veriler, Bilgi sistemleri, Bağlantılı açık veriler

* PhD student of the International PhD School for Humanities, SDISU (sdisu.unical.it/), University of Calabria, Italy, (iryna.solodovnik@gmail.com)

The Semantic Web technologies, Linked Data and SKOS

The Semantic Web* (or the intelligent Web, in reference to Web 3.0) is not a separate web, but an extension of the current web which contains the virtual realm of virtually boundless information in the form of web documents. Semantic web technologies are potential enough to make logical connections of and decisions (through inference rules) on different pieces of web data, fusing (Bleiholder et al., 2008) their meaning (semantics, ontology, “shared conceptualization”¹), and enabling computers and humans to work better in cooperation (Berners-Lee, 1998). Indeed, the Semantic Web amplifies the conceptualization of (meta)data² allowing it to become semantic entities responsible to organize, access, retrieve and preserve digital information resources. While “conceptually old for library and information professionals, metadata³ has taken a more significant and paramount role than ever before and is considered as the golden key for the next evolution of the web in the form of Semantic Web” (Safari, 2004, p.1; Karen, 2010). Semantic (meta)data contribute to semantic interoperability (Tolk et al., 2007) and cross-searching of web contents.

The Semantic Web itself requires adding semantic metadata on the top of (meta) data describing web resources. This approach is aiming at processing effectively the data based on the semantic information associated with data. In this way, computers can make inferences about the data, ‘understanding’ what data resources are and how they relate to other data.

“The Semantic Web provides a framework for making data more accessible and easier to harmonize. It has the potential to unlock information that would be difficult to uncover using traditional data technologies”⁴. It is only a matter of fact to get exponentially available data (already included in Web sites, other databases, XML documents, and other systems) into a uniform format such as RDF (Recourse Description Framework). Another step is to classify and to connect data considering its properties and its relationships with other data. This is also where Semantic Web technologies such as, RDFS (Recourse Description Framework Schema) and OWL (Ontology Web Language) come in.

* The term “Semantic Web” was conceived in 2001 by Tim Berners-Lee, the inventor of the World Wide Web.

- 1 Semantics of data on the web is often called ontology, that is a “formal, explicit specification of a shared conceptualization [to specify which] one needs to state axioms that do constrain the possible interpretations for the defined terms”. The shared conceptualization provides a shared vocabulary, which can be used to model a domain represented by different types of existing objects and/or concepts, and their properties and relations (Gruber, 1993). Today different ontologies, representing the third basic component of the Semantic Web, are used to formalize and enhance the semantic value of web data, included systems engineering, software engineering, biomedical informatics, library science, information architecture. Particularly, a program that wants to compare or combine information across the two databases must have a way to discover common meanings for whatever databases it encounters. A solution to this problem is provided by collections of information called, precisely, ontologies.
- 2 On the Web, the distinction between data and metadata is not absolute; sometimes the resource can be interpreted simultaneously in both modes, and metadata it itself data, which can be described by other metadata. Almost always, to avoid ambiguity, data and metadata are based on a specific syntax (logical structure).
- 3 METADATA: what in the world? <<http://pfs.parkhill.k12.mo.us/Users/11zhaoj/assignments/metadata.htm>>
- 4 LOD2, Creating Knowledge out of Interlinked Data, <<http://planet.lod2.eu/tag/linked%20data>>

Summarily, the Semantic Web, whose main purpose is to create a collaborative data infrastructure⁵ where to generate and exchange new knowledge, is aiming at:

- a. enriching information resources available in various forms on the Web through semantic annotations⁶ (Rusu et al., 2011), RDF crosswalks⁷ and formal descriptions/ontologies (or web vocabularies, taxonomies⁸ capturing semantics of metadata within schema structures)⁹ (Valkeapää et al., 2007);
- b. providing meaningful semantic connections by mechanisms of Linked Data (Heath et al., 2008; Baker, 2010), which is “the first practical expression of the Semantic Web”¹⁰. Linked Data infer new levels of knowledge within a global space of information resources, and also reveal new information behaviors and needs of different communities of users across the web. Links among different resources in the Linked Data Web will enable discovery of semantically related resources;
- c. creating a common agreed framework (through common ontology) of Cloud platforms¹¹. These platforms enable the exchange of information in a unified manner, and enhance the interaction of services and tasks performed by computers within a distributed network community (Kim et al., 2010; Mitchel, 2010). Through Cloud platforms, the consumer and producer agents can reach a shared understanding by exchanging ontologies, which provide the vocabulary needed for “discussion”.

The architecture of the Semantic Web provides the following basic technological components:

1. URI (Uniform Resource Identifier, “Globally Unique Identifier”¹²). The URI is a formalism used to identify uniquely an information resource on the web. Each resource (single

5 To get a quick idea about a collaborative data infrastructure (built on research data), see the report “A Surfboard for Riding the Wave Towards a four country action programme on research data”, <<http://www.knowledge-exchange.info/Default.aspx?ID=469>>

6 See TEXTUS, that has an extensible model for semantic annotations, <<http://textusproject.org/>>

7 Semantic Interoperability of Metadata and Information in unLike Environments (SIMILE) Project. MIT, 2008. “RDFizers-SIMILE” Cambridge Mass.: MIT, at <<http://simile.mit.edu/wiki/RDFizers>>

8 Taxonomy and a set of inference rules represent the most typical kind of ontology for the Web. “The taxonomy defines classes of objects and relations among them. For example, an address may be defined as a type of location, and city codes may be defined to apply only to locations, and so on. Classes, subclasses and relations among entities are a very powerful tool for Web use” (Berners-Lee et al., 2001).

9 Ontologies can improve the accuracy of Web searches: the search program can look for only those pages that refer to a precise concept instead of all the ones using ambiguous keywords. More sophisticated applications will use ontologies to relate the information on a page to the associated knowledge structures and inference rules. Here is an example of application for e-learning based on ontological structures: <<http://www.merlot.org/>> (if different materials are organized into units -learning objects - each unit can be connected to others and reassembled in a new course). For an overview of possible applications based on ontologies: <<http://www.netcrucible.com/semantic.html>>; Maedche (2002).

10 <<http://linkeddata.org/>>

11 The Cloud Computing Interoperability Forum, <<http://www.cloudforum.org/>>; Welcome to the Data Cloud? The Semantic Web, October 6, 2008, <<http://www.zdnet.com/blog/semantic-web/welcome-to-the-data-cloud-/205>>; Semantic Cloud Computing. Bringing Semantics to the Cloud, <<http://www.fluidops.com/semantic-cloud-computing/>>

12 “Linked data, libraries, and the Semantic Web”: <http://www.gordondunsire.com/pubs/pres/linkeddata/libsemweb.ppt>

document, its parts and its metadata¹³, objects and entities mentioned in them, image, service, e-mail) must have its URI, which can be a Web address (URL, Uniform Resource Locators, the most common type of URI) or a namespace (URN) (Berners-Lee et al., 2005). The Semantic Web, in naming every concept simply by a URI, allows anyone express new invented concepts with minimal effort.

2. RDF (Recourse Description Framework)¹⁴. The RDF is a W3C standard representing a declarative meta-language based on a XML-based model for encoding, exchange and reuse of metadata and their patters on the Semantic Web. The RDF provides itself the RDF data model based on the 3-part statement (triple): Resource (Subject), Property/relation (Predicate), Value (Object) (Figure 1). The automatic data triplification, according to this exposure, is closely related to human way of thinking and building concepts.

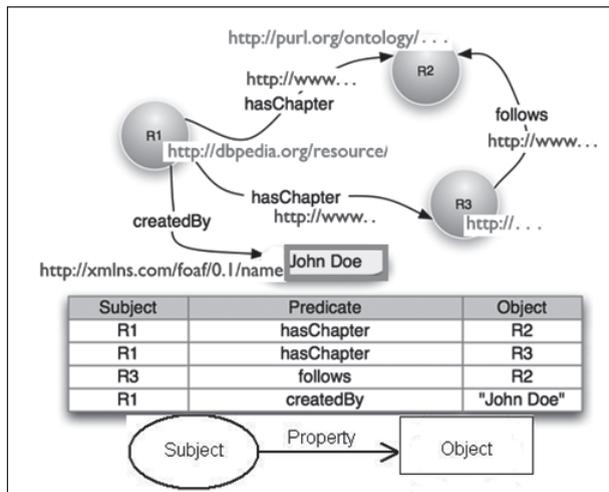


Figure 1. RDF graph of triples¹⁵

In Figure 1, the sets of linked triples are shown as a graph of nodes and connectors identified by URIs. It is possible to imagine RDF triples connecting different data like hypertext links connecting a set of document on the web. RDF triples, which can be written with XML tags, specify relationships between “Subjects” and “Objects”, in order

13 “Metadata may refer to any resource which has a URI. Metadata may be stored in any resource no matter to which resource it refers”(Berners-Lee, 1997)

14 <http://www.w3.org/RDF/>

15 Based on a source from Open Archives Initiative, Object Reuse and Exchange, CC3.0 License, <http://www.openarchives.org/ore/1.0/primer>

to navigate between them. This approach provides the integration of information from multiple resources, as well as allows the fluent automatic access to different related data despite their diversified origin.

The nodes "Object", represented by rectangular, contain data that can be both literal and URI. These nodes form terminators of linked data chains because they cannot be matched to other nodes without ambiguity" (Dunsire, 2012). Nodes of "Subjects" and "Predicates" are identified by URIs, and can be processed only by machines. Summarily, the unifying logical syntax of RDF triples enables different concepts defined by URIs to be progressively linked into a universal Web, and to make logical assertions based on the associations between "Subjects" and "Objects"; thus generating automatically web statements about resources. The inference among RDF predicates is possible through

3. *Inference engines* (web agents) or computer programs capable of interpreting RDF and OWL semantic information. They are essential component in the generation of new knowledge on the web. Indeed, the potential of the Semantic Web would be useless if there were no such inference engines gathering information from diverse sources, as well as processing and exchange it with other programs, inferring new data.

While RDF model provides a good syntax for web resources, it does not specify their semantics. For this reason, the Semantic Web offers the already cited technologies such as RDFS and OWL.

4. The RDFS¹⁶ represents a vocabulary describing groups of related RDF resources together with their relationships. Particularly, an RDFS vocabulary expresses the acceptable properties and their values that may be assigned to RDF resources within a given domain. Moreover, RDFS' mechanisms permit to create classes of resources (that become instances of classes) sharing common properties as well as relationships among these resources. In their turn, classes are resources too, and any class may represent a subclass of another. This hierarchical semantic information structure is what allows computers to determine the semantics of resources based on their properties and classes.
5. The OWL¹⁷, built upon RDFS, is the richest standard description web vocabulary available today to define web ontologies used to create advanced Semantic Web applications (O'Connor et al., 2008). These ontologies consist of a taxonomy (system of classification)¹⁸ and a set of inference rules from which automatic logical deductions (conclusions) can be made. The OWL syntax (e.g. *subClassOf*, *disjointWith*,

16 <<http://www.w3.org/TR/rdf-schema>>

17 <<http://www.w3.org/TR/owl-features/>>

18 Taxonomy is a system grouping resources into classes and sub-classes based on their relationships and shared properties.

unionOf, intersectionOf) allows to assign properties to classes of resources as well as permits their subclasses to inherit the same properties. The growing expressive complexity of OWL finds its fittings in three sub languages, such as OWL Lite, OWL DL, and OWL Full¹⁹ (Lacy, 2005), each with enhancing level of detail required by different web semantic models.

6. Statements built on RDF structures are queried by means of *SPARQL* (Simple Protocol and RDF Query Language)²⁰.

Here below (Figure 2) there is a graphical representation of the Semantic Web technologies.

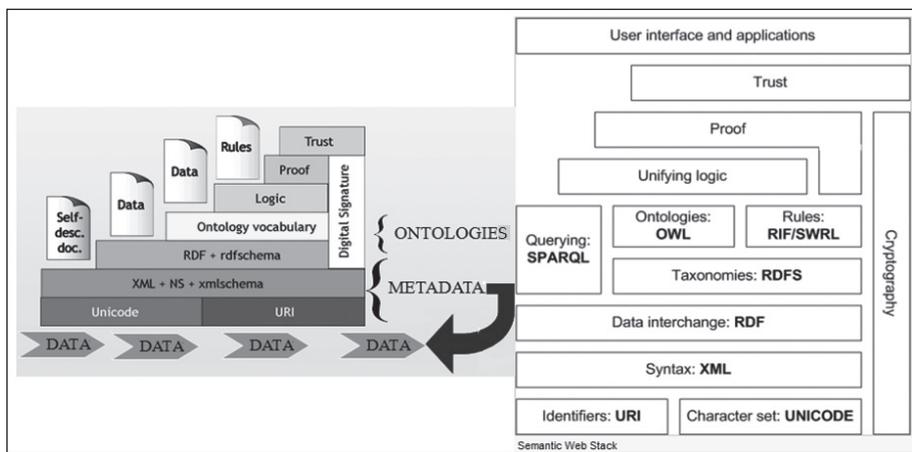


Figure 2. The Semantic Web Technologies Stack²¹

As we can see from Figure 2, over the ontological level of the Semantic Web stack, there is the Logic level managed by SWRL²² language. This level should provide automated reasoning and inference of machine understandable knowledge, allowing it to be automatically integrated and reused by web applications. To achieve the full

19 Difference between OWL Lite, DL, and Full, available at Memento, 2007, <<http://ragrawal.wordpress.com/2007/02/20/difference-between-owl-lite-dl-and-full/>>. OWL 2 Web Ontology Language Document Overview, W3C Recommendation 27 October 2009, <<http://www.w3.org/TR/owl2-overview/>>; OWL 2 DL ontologies for terms of the Dublin Core Metadata Initiative. It is meant for applications and other ontologies which need OWL DL versions for reasoning or import rather than the existing RDFS schemas provided by the Dublin Core Metadata Initiative itself, <http://bloody-byte.net/rdf/dc_owl2dl/index.html>

20 SPARQL query language bears close resemblance to SQL only applicable to an RDF data graph, <<http://www.w3.org/TR/rdf-sparql-query/>>

21 The Semantic Web Architecture, <<http://obitko.com/tutorials/ontologies-semantic-web/semantic-web-architecture.html>>; http://en.wikipedia.org/wiki/Semantic_Web#cite_note-16>

22 SWRL (<<http://www.w3.org/Submission/SWRL/#2.1>>) is proposal for a semantic web rules-language, combining sublanguage of the OWL Web Ontology Language (OWL DL and Lite) with those of rule markup language (RuleML).

potential of the Semantic Web, information must be approved at the Proof level, permitting humans to retrace the steps a Semantic Web agent took to arrive at a particular conclusion. Finally, the entire Web would benefit of reliability and security (Weitzner et al., 2007) of web information validated through digital signatures at the Trust level²³. Digital signatures are encrypted blocks of data that computers and agents can use to verify that the attached information has been provided by a specific trusted source. As Miller (2009) put it in one podcast, “the Semantic Web will expose all of the problems of the Web like trust, provenance, and reliability (problems which are already very much with us) in a large distributed space”.

One of the fundamental problems of the Semantic Web is to make available various types of data, so that they could be integrated and interoperable. Technically, it can be achieved through the appropriate technologies converting different datasets formats into RDF. “The process of converting all existing data to RDF can be a major hurdle for organizations with large numbers of unstructured text documents and few metadata experts. Many tools have been developed to help automate named entity recognition, which is the process of using software to automatically identify and classify text elements like the names of persons, organizations, geographical locations, expressions of time, or expressions of quantity” (Goddard et al., 2009).

The available technologies such as, for example, POWDER²⁴, RDFa²⁵ GRDDL²⁶, R2RML²⁷, RIF (Kifer, 2008), Drupal7 (Corlosquet et al., 2010) allow to make an automatic RDF mark-up of websites. To transform structural data in RDF/OWL formats there are tools such as *Web services links & resources*²⁸; *SemWev*²⁹; *Beckett*³⁰, *SIMILE* (RDF crosswalks)³¹, *Semantic Bank* (Huynh, et al., 2005), *D2R Server* (exposure of related databases in RDF) (Bizer et al., 2009), *Altova SemanticWorks* (the ground-breaking visual RDF/OWL editor)³². Moreover, to convert unstructured text into blocks of main entities, topics, reports; as well as to perform extraction of keywords, auto-tagging and disambiguation of entities and concepts - which may serve as outputs for RDF - there are several semantic tagging tools API (*Application Programming Interface*) such as, for example, *OpenCalais*³³ e *Zemanta*³⁴.

23 The Semantic Web: Proof, Trust, And Security by T.Welsh, Editor, Web Services Strategies, available from Cutter Consortium's bookstore, 23 Sept. 2003, <<http://www.cutter.com/research/2003/edge030923.html>>

24 <<http://www.w3.org/TR/powder-dr/>>

25 <<http://en.wikipedia.org/wiki/RDFa>>

26 <<http://www.w3.org/TR/grddl-primer/>>; GRDDL Use Cases: Scenarios of extracting RDF data from XML documents, <<http://www.w3.org/TR/grddl-scenarios/>>

27 <<http://www.w3.org/TR/r2rml/>>

28 <http://www.wsindex.org/Companies/Semantic_Web/index.html>

29 <http://semanticweb.org/wiki/Main_Page>

30 <<http://planetrdf.com/guide/>>

31 RDFizers — SIMILE, <<http://simile.mit.edu/wiki/RDFizers>>

32 As an example of automatic links creation, see Google-Refine tool, a power tool for working with messy data, cleaning it up, transforming it from one format into another, extending it with web services, and linking it to databases like Freebase, <http://www.altova.com/download/semanticworks/semantic_web_rdf_owl_editor.html>

33 <<http://www.opencalais.com/calaisAPI>>

34 <<http://developer.zemanta.com/>>

Despite the existence of different tools for automatic RDF metadata publishing (Berrueta et al., 2008), included for semantic link discovery (Hassanzadeh, 2009; Volz et al., 2009), the development of links³⁵ between different datasets is not a trivial process, because it is necessary to carefully calculate an organic re-use of data shared within different user communities. It is also worth noting that the process of creating links may be carried out both through manual and ad hoc algorithms, expressing explicitly properties and values of (meta)data and constraints imposed on them. Anyway, considering how different technologies have been proliferating to support the creation of RDF/OWL, it is likely that the Semantic Web vision is one that will be realized globally in the near future.

“Exposing data as RDF is an important first step, but to actually achieve the linked-data vision we must set explicit RDF links between data items within different data sources. This provides the means by which we can discover more information about a given entity” (Goddard et al., 2009). To actually link the Semantic Web datasets, in 2006 Tim Berners-Lee - in his memorable web document “Linked Data”³⁶ - proposed a new Semantic Web technology called Linked Data (LD). LD is a technology based on:

1. RDF (to provide useful information on the object),
2. Hyper Text Transfer Protocol/HTTP (so that these objects can be referenced, searched and accessed by user agents), and
3. dereferenceable URIs identifying objects, “emphasizing data interconnections, interrelationships and context useful to both humans and machine agents”³⁷.

By means of these supporting tools, LD provides best practices for publishing (Bizer et al., 2007), exposing, connecting and sharing different data(sets)³⁸ across the web. The main purpose of LD is to break down the technological barriers that prevent free data sharing (Bizer et al., 2009), and to enable more powerful exploration of linked datasets (Alexander et al., 2009) structures through SPARQL queries (Cyganiak et al., 2008). However, according to some authors (Bechhofer, S., Ainsworth, J., Bhagat, J., Buchan, I., Couch, P. and Cruickshank, D., 2010), LD still misses a mechanism describing the aggregation of resources making their relations well interpretable, in order to capture better the added value of data collections and to allow its reuse through the exchange of a single object.

35 Google Refine is a power tool for working with messy data, cleaning it up, transforming it from one format into another, extending it with web services, and linking it to databases like Freebase.

36 Tim Berners-Lee, Linked Data (Editing status: imperfect but published), last change, 2009, <http://www.w3.org/DesignIssues/LinkedData>; Berners-Lee promoted further LD during the TED conference, 2009, http://www.ted.com/talks/tim_berniers_lee_on_the_next_web.html

37 Linked Data FAQ, http://structureddynamics.com/linked_data.html#question_9

38 Datasets may derive from relational databases; from interoperable and non-interoperable information repositories as, for example, Electronic data interchange (EDI) systems used for the structured transmission of data between organizations by electronic means; from XML documents, and other systems increases the amount of useful data available exponentially.

In Figure 3 there is provided a graphic representation of LD datasets³⁹ published on the web.

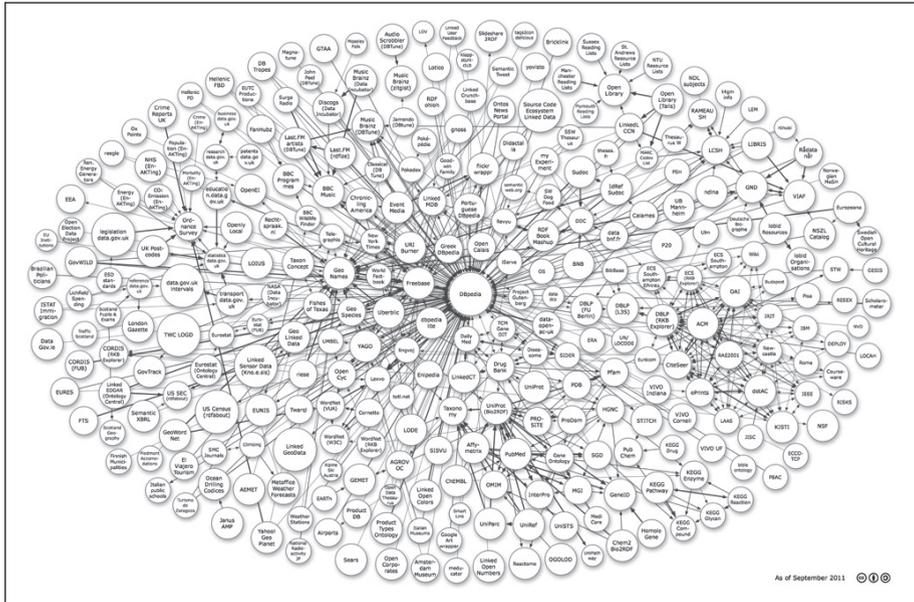


Figure 3. Connecting Web Dataset Through *Linked Data*

Source: "Linking Open Data cloud diagram", <<http://richard.cyaniak.de/2007/10/loj/>> (clicking the original image will take you to an image map, where each dataset is a hyperlink to its homepage)

The graph showed on Figure 3 is a result of efforts of the Linking Open Data community project⁴⁰, within the W3C SWEOW⁴¹ group. The project uses the categories of datasets converging to a directory of Open Data datasets and Linked Open Data (LOD) called CKAN⁴² managed by the Open Knowledge Foundation⁴³. While in October 2007 the datasets of LOD cloud diagram were based on more than 2 billion RDF triples connected by more than 2 million RDF links (Berners-Lee, 2009); in 2011 the datasets counted 31 billion RDF triples connected by around 504 million RDF links. To determine whether LD technologies are sufficiently mature, there can be explored development and deployment of exposing data as RDF, and linking RDF entities together.

39 Here there are data sets that are available on the Web as LD and contain data links pointing at other LD sets, <<http://thedatahub.org/group/lojcloud>>

40 SweoIG/TaskForces/CommunityProjects/LinkingOpenData <<http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>>

41 LinkingOpenData, W3C SWEOW Community Project, <<http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>>

42 <<http://ckan.net>>

43 Open Knowledge Foundation, <<http://okfn.org/>>

Referring to the practical use of LD datasets, imagine a system such as, for example, an Institutional repository (IR). Here many contributions (articles, books and their parts, thesis, conference proceedings) containing good bibliographical indications (including those of websites) may be deposited. The adoption of linking mechanisms could enrich and enhance these indications through connecting (Hennekenet al., 2011; Woutersen-Windhouver, 2009) them with other citations, entries of encyclopedias, glossaries, classifications and other value (controlled, authority) vocabularies published on the web as LD structures.

The relative abundance of potential links can be, in their turn, enriched with thousands of other links pointing to another information resources, registers of agents (people) and their curricula, organizations, generating a powerful cross-border flow of information and data. "Making entities identifiable and referenceable using URIs augmented by semantic, scientifically relevant annotations greatly facilitates access and retrieval for data which used to be hardly accessible"⁴⁴. Really, LD publishing, sharing and interlinking scientific resources and data is intended to extend and fully realize the potential of access to scientific resources and collaboration within and across disciplines, whose knowledge is exposed and conveyed on the Web (Heath et al., 2011).

To control the quality of data exposed through LD mechanisms, there is a need to validate it by means of authority data inherent to Knowledge Organization Systems (KOSs)⁴⁵. KOSs consist of authority systems such as thesauri⁴⁶, classification schemes, subject heading lists, taxonomies and others controlled vocabularies. To port the already existing KOSs on the web (Tudhope, 2004; Zeng, 2009), as well as to provide conceptual modeling language for developing and sharing new KOSs, the W3C has developed the Simple Knowledge Organization System (SKOS). Particularly, SKOS is an application of the RDF⁴⁷ and its details have been released in "SKOS Reference"⁴⁸ together with a user guide "SKOS Primer"⁴⁹. SKOS is aimed at building a bridge between KOSs (used in libraries, archives, museums, government portals, enterprises, social network applications and other communities) and LD community, bringing benefits to both. Indeed, nowadays, "Libraries, museums, newspapers, government portals,

44 <<http://linkedscience.org/events/lisc2011/>>

45 Knowledge Organization Systems: An Overview, Council of Library and Information Resources, <<http://www.clir.org/pubs/reports/pub91/1knowledge.html/#1>>. See also Networked Knowledge Organization Systems (NKOS) Registry Reference Document for Data Elements – Draft. Last formatted: August 20, 2008, <<http://nkos.slis.kent.edu/registry3.htm>>. From the NKOS registry there can be selected values for the expansion of terms through thesauri and other controlled vocabularies, classification schemes, usage notes, conceptual relationships, data entry, spelling variants.

46 The new standard for Thesauri ISO 25964:2011 "Information and documentation. Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval" replaced ISO:2788 and ISO:5964, New international thesaurus standard Published (Press Release), <<http://nkos.slis.kent.edu/PressReleaseISO25964-1Dec2011.pdf>>

47 SKOS is an application of RDF-PRIMER (<<http://www.w3.org/TR/rdf-primer/>>), which is the instance of OWL-SEMANTICS (<<http://www.w3.org/TR/2008/WD-skos-reference-20080829/skos.html#OWL-REFERENCE>>)

48 <<http://www.w3.org/TR/2009/PR-skos-reference-20090615/>>

49 SKOS Primer, 2009, <<http://www.w3.org/TR/2009/NOTE-skos-primer-20090818/>>

enterprises, social networking applications, and other communities that manage large collections of books, historical artifacts, news reports, business glossaries, blog entries, and other items can use SKOS to leverage the power of Linked Data⁵⁰.

SKOS represents an excellent way for conceptual exposure, management, sharing and re-use of authority data on the web, linking them with related authority data and integrating with different metadata conceptual schemes [e.g. with Dublin Core⁵¹, Library of Congress Subject Headings, MARCXML (Summers et al., 2008)]. This is possible by identifying concepts provided by KOSs with URIs, labeling them with strings in one or more natural languages, as well as documenting them with different types of note, semantically related to each other, and aggregated into concept schemes. The growing scenario of use cases⁵² implementing SKOS offers the prospect of linking together vocabularies provided by different sectors, thus enhancing "authority control"⁵³ of their data exposed on the web. The task to validate data through authority data ensures compatibility between different datasets, as well as their harmonized automatic management and interoperability at an aggregate level. "Finally, the SKOS vocabulary itself can be extended to suit the needs of particular communities of practice or combined with other modeling vocabularies"⁵⁴.

A knowledge global network sharing datasets outputs enhanced with LD and SKOS approaches is a perspective of a few years, considering also that there are enough good practices and use cases to be imitated⁵⁵.

Linked Open Data and Supporting Experiences

The LOD cloud diagram, already presented, assumes that different datasets "must be provided in such a form that there are no technological obstacles to share data. This can be achieved by the provision of the work in an Open Data format, i.e. one whose

50 Using SKOS to leverage the power of Linked Data, <<http://www.w3.org/2004/02/skos/>>

51 Data Catalog Vocabulary/DC-SKOS, <http://www.w3.org/egov/wiki/Data_Catalog_Vocabulary/DC-SKOS>

52 SKOS Use Cases and Requirements, 2009, <<http://www.w3.org/TR/2009/NOTE-skos-ucr-20090818/>>. Usage examples: Publishing 20th Century Press Archives, Data BNF, LOCAH, Browsing And Searching In Repositories With Different Thesauri, Component Vocabularies, PODE, Subject Search, Europeana, VIAF, AGROVOC Thesaurus, AGRIS, Vocabulary Merging (SKOS mapping), Migrating Library Legacy Data, NLL Digitized Map Archive, Collecting material related to courses at The Open University. See also Thacker, M. SKOS and URIs, as used in the Local Government Business Model, Standard Hub, <<http://standards.data.gov.uk/proposal/skos-and-uris-used-local-government-business-model>>

53 Lanius, L., Implementing Authority Control. An online workshop offered by the Vermont Department of Libraries, <<http://libraries.vermont.gov/sites/libraries/files/tsu/implementingauthoritycontrol.htm>>

54 <<http://www.w3.org/TR/2009/NOTE-skos-primer-20090818/>>

55 LinkedScience.org, <<http://linkedsience.org/tag/linked-data/>>. The Digital Curation Centre (<<http://www.dcc.ac.uk/>>) published Cite Datasets and Link to Publications (<<http://www.dcc.ac.uk/resources/how-guides/cite-datasets/>>), the guide that illustrates how to create links between research publications and data on which they are based, thereby making it possible to locate the dataset for those who read an article, and vice versa. See also: LOD2, Create Knowledge out of interlinked data, <<http://lod2.eu/Welcome.html>>; Application of Linked Data for Authority data enrichment, <http://www.w3.org/2005/Incubator/ld/wiki/Use_Case_Authority_Data_Enrichment>; Semantic alignment: expressing Library Data through existing Linked Data vocabularies, <http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/#Semantic_alignment>; Vocabulary alignment <http://www.w3.org/2005/Incubator/ld/XGR-ld-usecase-20111025/#Use_Case_Authority_Data_Enrichment>

specification is publicly and freely available and which places no restrictions monetary or otherwise upon its use⁵⁶. This means, that the LD paradigm matches well with the vision of Open Data.

According to Open data definition, Open data, is a piece of open knowledge that is “free to use, reuse, and redistribute”⁵⁷ under an unrestricted license such as, for example, Creative Commons (CC) licenses, and Talis⁵⁸ (Campbell et al., 2010). The goals of the Open Data are inherent in several Open movements, such as Open access⁵⁹, Open content⁶⁰, Linked open science⁶¹, Open Knowledge⁶², Open Government⁶³, Open Bibliographic Data⁶⁴, Open Source (Buitenhuis et al., 2010) and others.

Open Data together with Linked Data⁶⁵ aim to break down the social, cultural, legal and economic barriers freeing shared data between human and software agents. Newly released Open Data published with LD mechanisms may be directly linked to the already existing open datasets (e.g. DBPedia.org, Wikipedia, WikiGuida, Geonames, MusicBrainz, lexical ontology WordNet, DBLP bibliography) exposed within the LOD cloud, thus reducing duplication of data and, above all, keeping the data updated⁶⁶, and offering different agents to discover new information and to create and share freely new knowledge.

As the Open Data and LD paradigms are developing into a mainstream topic, more and more organizations are announcing new projects and services making their data open and publishing it as LD. Furthermore during the last two years Open Data (Danowski, 2010) and LOD has received a lot more attention from the library world (Dunsire, 2012). Here are some examples of related, practical experiences:

- ◇ The Harvard Library Policy on Open Metadata is committed to providing “Open Access to library metadata, subject to legal and privacy factors. In particular, the Library makes available its own catalog metadata under appropriate broad use licenses”⁶⁷.

56 Open Definition, OKD, <<http://opendefinition.org/okd/>>

57 OKD, <<http://www.opendefinition.org/>>

58 <<http://creativecommons.org/>>; <<http://www.talis.com/tdn/tcl/>>

59 Suber, P. Open Access Overview. Last revised March 18, 2012, <<http://www.earlham.edu/~peters/fos/overview.htm>>

60 <<http://opencontent.org/definition/>>

61 <<http://linkedsience.org/tag/linked-open-science/>>

62 Open Knowledge Foundation. Promoting Open Knowledge in a Digital Age, <http://okfn.org/>. Dell'attività del gruppo di lavoro sui dati aperti della Open Knowledge Foundation si parla nell'articolo di Jennifer C. Molloy "The Open Knowledge Foundation: Open Data Means Better Science",

63 <<http://www.data.gov/>>

64 < <http://obd.jisc.ac.uk/>>

65 "While, to date, it is the case that linked data has been demonstrated using public Web data and many desire to expose more through the open data movement, there is nothing preventing private, proprietary or subscription data from being Linked Data [...] Since linked data can be applied to any data formalism, source or schema, it is perfectly suited to integrating data from inside and outside the firewall, open or private", Linked Data FAQ, <http://structuredynamics.com/linked_data.html#question_9>

66 Tim Berners-Lee Talk on TED, 2009, <<http://www.ted.com/pages/about>>

67 <<http://openmetadata.lib.harvard.edu/>>

- ◇ The Library of Congress has developed the 'Authorities and Vocabularies'⁶⁸ service exposing its terminological systems and standards in an open manner and publishing them as LD.
- ◇ The German National Library (DNB) has developed 'Authority Data Linking' connecting its bibliographic data with Wikipedia, DBpedia, and VIAF datasets (Keßler, 2010). In cooperation with German Serials Database (Zeitschriftendatenbank - ZDB)⁶⁹, the DNB has also generated Linked Data Service⁷⁰. This service publishes DNB's bibliographic data and authority data as LOD, under the CC0 license, making such data available according to Open definition⁷¹. The record structures expressed in RDF/XML are available on the portal DNB⁷² representing an experimental service which is going to be continually expanded and improved in accordance with the transparent procedures of the public domain⁷³.
- ◇ The Hungarian National Library has published its bibliographic and authority data in open modality, using RDFDC, FOAF, and SKOS⁷⁴.
- ◇ The British National Library (BNB) has provided its data via RDF download⁷⁵. This practice of converting data is going beyond the encoding of collections of MARC records (over 2.8 million) in RDF/XML. Moreover, the BNB is working in the direction of making available the British National Bibliography (BNB)⁷⁶ as LOD through Talis platform⁷⁷, connecting its data with LOD datasets such as VIAF, LCSH, Lexvo, GeoNames, MARC country, Dewey.info, RDF Book Mashup⁷⁸.
- ◇ The *Bibliothèque Nationale de France* is carrying out the *Data.bnf.fr* project, which is aiming to elaborate the bibliographic data (authors, works) in RDF triples publishing them in LOD modality⁷⁹.
- ◇ The Ontology Engineering Group (OEG) of the *Universidad Politécnica de Madrid* (UPM) has launched an open project "Linked data at the BNE" (*Biblioteca Nacional de*

68 Library of Congress Authorities, <<http://authorities.loc.gov/>>; Authorities&Vocabularies, <<http://id.loc.gov/authorities/about.html>>

69 <<http://www.zeitschriftendatenbank.de>>

70 <<https://wiki.dnb.de/display/LDS/>>

71 <<http://creativecommons.org/publicdomain/zero/1.0/>>; <<http://opendefinition.org/>>

72 <<http://portal.dnb.de/>>

73 For more information: <<http://openbiblio.net/2012/01/26/german-national-library-goes-lod-publishes-national-bibliography/>>; <http://files.d-nb.de/pdf/linked_data.pdf>; <http://files.d-nb.de/pdf/linked_data_e.pdf>

74 Hungarian National Library OPAC and Digital Library Published as LD. ISKO UK, 2010, <<http://iskouk.blogspot.com/2010/05/hungarian-national-library-opac-and.html>>

75 Wilson, N. Linked Data Prototyping at the British Library. British Library, 2010, <<http://talis-linkeddata-libraries.s3.amazonaws.com/Linked%20Data%20Prototyping.pdf>>

76 <<http://www.bl.uk/bibliographic/natbib.html>>; <http://semanticweb.com/british-library-announces-major-release-of-linked-data_b21499>

77 <<http://www.talis.com/platform/>>

78 Wilson, N. Linked Data Prototyping at the British Library. British Library, 2010, <<http://talis-linkeddata-libraries.s3.amazonaws.com/Linked%20Data%20Prototyping.pdf>>; <<http://thedatahub.org/dataset/bluk-bnb>>

79 Data.bnf.fr: <<http://data.bnf.fr/>>; <<http://data.bnf.fr/docs/databnf-presentation-en.pdf>>; <<http://thedatahub.org/it/dataset/data-bnf-fr>>; <<http://openbiblio.net/2011/09/21/lod-at-bibliotheque-nationale-de-france>>

España)⁸⁰. The project has developed the tool MARiMbA⁸¹, by means of which BNE's bibliographic data are getting connected with LOD data source for fields appropriate for authority control. The MARiMbA tool generates RDF metadata from MARC21, using RDFS/OWL vocabularies which can be queried through SPARQL⁸². Besides the converting of MARC21 in RDF, the BNE has started to publish its bibliographic and authority data directly in RDF, according to the principles of LD and open CCO licenses.

- ◇ The University of Münster (Germany) has launched the project "LODUM" (Linked Open Data University of Münster)⁸³ based on the Open Access strategies⁸⁴. The LODUM is aimed at integrating in the LODUM infrastructure scientific data and publications, other data (class schedules, administrative data) of the same university, releasing them as LOD.
- ◇ The project *alfa Linkypedia*⁸⁵ collects all certified links supporting items of Wikipedia. Since many of harvested links contain a variety of citations⁸⁶ from library, museum and archive domains, the *Linkypedia* can surely be useful to support open citation practices within innovative information management systems on the web.
- ◇ The *Bayerische Staats Bibliothek*, together with other libraries, is taking steps to align⁸⁷ its (meta)data with Europeana Data Model (EDM) (Doerr, M. et al, 2010) (Figure 4).

The EDM model, showed in Figure 4, enables to reuse the RDF, RDFS, OAI-ORE⁸⁸, SKOS, DCMI Terms⁸⁹ namespaces (describing digital bibliographical records), outputting bibliographic data in LOD. Europeana's professional knowledge sharing platform⁹⁰, based on this model, is a multi-lingual online collection of millions of digitized items derived from European museums, libraries, archives and multi-media collections. Europeana projects converts different terminologies and other KOSs (provided by various cultural

80 <<http://www.bne.es/es/Catalogos/DatosEnlazados/index.html>>

81 2.4 million bibliographic records (ancient and modern monographs, sound recordings) and 4 million authority records (personal names, organizations, uniform titles and subjects) are converted into RDF about The transformation process has generated approximately 58 million RDF triples and 600 links (owl:sameAs) enriching datasets such as DBpedia and VIAF, <<http://thedatahub.org/dataset/datos-bne-es>; <http://mayor2.dia.fi.upm.es/oeg-upm/index.php/en/downloads/228-marimba>; <http://thedatahub.org/dataset/datos-bne-es>>

82 <<http://datos.bne.es/sparql>>

83 Linked Open Data University of Münster, <<http://lodum.de/about>>; LinkedScience.org, Analyzing and Visualizing Productivity of a University, <<http://linkedsience.org/tag/lodum/>>

84 <<http://unesdoc.unesco.org/images/0021/002158/215863e.pdf>>

85 Linkypedia, <<http://linkypedia.inkdroid.org/>>

86 Cite Datasets and Link to Publications, published by the Digital Curation Centre: <<http://www.dcc.ac.uk/>>, <<http://www.dcc.ac.uk/resources/how-guides/cite-datasets>>

87 Europeana Libraries: Aggregating digital content from Europe's libraries. Report on the alignment of library metadata with the EDM, <<http://www.europeana-libraries.eu/documents/868553/1eade085-34ac-487f-82afd5cd2545e619>>; <<http://thedatahub.org/group/bibliographic>>; <<http://pro.europeana.eu/linked-open-data>>

88 <<http://www.openarchives.org/ore/>>

89 <<http://dublincore.org/documents/dcmi-terms/>>

90 <<http://www.europeana.eu/portal/>>

institutions)⁹¹ in SKOS, publishing them as LOD⁹². To support Europeana's initiatives, also the "European Commission (EC) is to fund the development of Linked Data tools that will enable more libraries and archives to provide digital content to Europeana [and to] enable innovative re-use of Europeana data in teaching and research contexts"⁹³.



Figure 4. The Namespaces Used in the Europeana Data Model (EDM) Model

Other experiences and initiatives ranging from:

- ◇ The knowledge-sharing platform "LinkedScience.org: Interconnecting scientific assets"⁹⁴;
- ◇ The Library of Congress' initiative "A Bibliographic Framework for the Digital Age"⁹⁵;
- ◇ The recently developed Bavaria's Open data portal⁹⁶;
- ◇ Sweden's national library system LIBRIS⁹⁷ publishing the Swedish National Bibliography along with the authority data under an open CC0 license provided as a complement to their LD implementation;

91 The "Data Exchange Agreement" (2012), <<http://pro.europeana.eu/documents/858566/c0c6e31f-5174-4898-9771-f9b9a8d1d4d7>>

92 Europeana LOD (the data.europeana.eu pilot) is part of Europeana's ongoing effort of making its metadata available as Linked Open Data on the Web. It allows others to access metadata collected from Europeana providers, via standard Web technologies, enrich this metadata and give this enriched metadata back to the providers, <<http://version1.europeana.eu/web/lo/d/datasets>>; <<http://pro.europeana.eu/linked-open-data>>; Europeana LOD Pilot datasets has already released about 2.5 million of records under the CC0 license.

93 <http://pro.europeana.eu/about?utm_source=portalmenu&utm_medium=portal&utm_campaign=Portal%2Bmenu>

94 <<http://linkedsience.org/>>

95 Today it's not enough for a library to be able to store collections. The space itself has to be engaging and inspiring to facilitate the users need for interoperable information, experiences and cultural inspiration, <<http://www.loc.gov/marc/transition/news/framework-103111.html>>

96 <<http://epsiplatform.eu/content/bavaria-opens-data-portal>>

97 <<http://librisbloggen.kb.se/2011/09/21/swedish-national-bibliography-and-authority-data-released-with-open-license/>>

- ◇ The Conference of European National Librarians (CELN)⁹⁸ and their vote to support the Open licensing of their data to groups like LOD-LAM (Linked Open Data in Libraries, Archives, and Museums)⁹⁹ (Oomen et al., 2012), IFLA'S Semantic Web Special Interest Group (SWSIG)¹⁰⁰, to library system vendors and providers, discussing and experimenting with LD technology, clearly reflect that LOD has become gained a lot of impetus in science, library and other cultural domains. The question is how to ensure that LOD won't be a temporary hype but that it will take hold in future infrastructures, generating LOD datasets from legacy systems and promoting the LOD approach towards a global and open information space.

To support the promotion of Open Data, different communities make their efforts to elaborate principles for releasing Open Data. Thus, the Open Bibliographic Data Working Group of the Open Knowledge Foundation has recently published "Principles for Open Bibliographic Data"¹⁰¹.

The concept 'bibliographic data' refers to data [e.g. author(s), title, publisher, date, title, page information, format of work] describing bibliographic resource as a unique resource in the set of all bibliographic resources, indicating also the modality how to find [e.g. URL address; URI identification: URN, DOI; ISBN, LCCN, OCLC number; links to related content, etc.] the described resource.

Formally, the Open Bibliographic Data Working Group recommends to release open bibliographic data or its sets with clear and explicit license¹⁰² statements regarding re-use and re-purposing of bibliographic elements. The licenses such as Creative Commons licenses (apart from CC0), GFDL, GPL, BSD (with non-commercial and other restrictive clauses) are considered not appropriate to release Open Data, because they hinder to effectively integrate and re-purpose datasets, preventing also commercial activities that could be used to support data preservation. The "Principles" establish that open bibliographic data should be explicitly placed in the public domain via the use of the Public Domain Dedication and License (PDDL) or CC0

Developing recommendations, upcoming challenges, and technical alignment of catalogues and legacy systems in cultural institutions, and authoring environments for scholarly communication with open data and service infrastructure based on Semantic Web principles will be the strategic ones for practical promotion of LOD approaches. Moreover, the related experiences needs to be supported not only by "education" of interested parties regarding the use of correct licenses and LD techniques, but rather

98 <https://app.e2ma.net/app/view/CampaignPublic/id:1403149.7214447972/rid:48e64615892ac6adde9a4066e88c736c>

99 LOD-LAM website (<<http://lod-lam.net/summit/>>) has grown into an active knowledge-sharing platform.

100 SWSIG, <<http://www.ifla.org/en/about-swsig>>; <<http://www.ifla.org/en/swsig>>; <<http://www.ifla.org/en/rss/group/6155>>

101 <<http://openbiblio.net/2010/10/15/principles-for-open-bibliographic-data/>>

102 <<http://opendefinition.org/licenses/#Data>>

concerning a change or a transformation of the “mental architectures” of sharing data, facts, information.

Other Use Cases of Linking Data at the Semantic Level

Making explicit links among different data, especially those at the semantic level, requires careful analysis and rigorous definition of all necessary features of a (meta) data system. As it was already mentioned, this can be achieved through the definition of a formal explicit and shareable specification (ontology) identifying concepts, their properties, values and relationships defining granularity of knowledge of a reference domain (Valkeapää et al, 2007). To link data of different knowledge domains exposed on the web, it would be a good practice to establish a common ontology (Sure, 2005; Vockner, 2011) for data sharing based on the already existing and widely used ontology structures¹⁰³. Establishing a common ontology will make data interpretable in a shared manner, thus also helping to create highly interoperable semantic web applications and services¹⁰⁴ (Sanfilippo et al., 2003; D’Aquin et al., 2008).

To show how data could be linked by means of common ontologies, here will be worth mentioning a cross-institutional project called “ResearchSpace”¹⁰⁵ carried out by the British Museum. This project aims at harmonizing data provided by different cultural organizations, using RDF to set up mechanisms for the semantic search¹⁰⁶. Particularly, the project uses a high level ontology aimed to improve search accuracy by understanding searcher’s intent and the contextual meaning of terms as they appear in the searchable dataspace. This ontology is based on the Conceptual Reference Model (CIDOC-CRM)¹⁰⁷ representing a data framework mapping links among different thesauri¹⁰⁸ terms supplied by “ResearchSpace”¹⁰⁹ users. The purpose of this methodology is to allow structured semantic search across multiple heritage repositories connected to GeoNames¹¹⁰ exposed LOD cloud

Joining the LD movement, in 2011 Google, Yahoo and Microsoft have agreed to adopt a common ontology maintained by Schema.org¹¹¹. This ontology permits to

103 An ontology is valid only for the domain for which it was designed.

104 <<http://www.w3.org/2009/Talks/1214-Darmstadt-IH/Applications.pdf>>

105 <www.researchspace.org>

106 Rather than using ranking algorithms such as Google’s PageRank to predict relevancy, Semantic Search uses semanticsto produce highly relevant search results.

107 <<http://www.cidoc-crm.org/>>

108 Thesauri are used both to control data entry, and to allow narrower-term searching of data, enabling to retrieve correctly information using synonym or near-synonym search terms.

109 <<http://poolparty.biz/dominic-oldman-skos-is-the-obvious-choice-for-representing-our-thesauri-in-semantic-form/>>

110 GeoNames is a geographical data base available and accessible through various Web services, under a Creative Commons attribution license, <<http://poolparty.biz/resources/glossary/item/?uri=http%3A%2F%2Fvocabulary.semantic-web.at%2FPoolPartySemanticWeb%2FGeonames>>

111 <www.schema.org>

publish linked structured data on the web, allowing different applications to create intelligent services systems (e.g. UMBEL Web Services, Virtuoso Universal Server, Linked Open Data Around-The-Clock)¹¹². Recently, the already mentioned “Europeana” has announced its new project called Europeana Libraries¹¹³, which aims to integrate into one research area digital collections of the best digital libraries from 11 European countries. The thematically categorized collections will be linked to the Google Books¹¹⁴ and to other web collections of photographs, manuscripts, historical films, to PhD thesis harvested by DART-Europe¹¹⁵, as well as to scholar articles from DOAJ (Directory of Open Access Journals)¹¹⁶. The Europeana Libraries collaborative platform will enrich data through a common ontology matching the ontologies of European libraries, and increasing the retrievability and re-use of their collections. In the light of the DRIVER (Digital Repository Infrastructure Vision for European Research) project, there was created an ontology-driven platform for semantic annotations to the ‘Academic Institutional Repository and Bibliography’¹¹⁷. The project has led to the discovery of new ways of semantic data exchange and, consequently, to the improvement of harvesting of the semantically related data. In order to link different disciplinary fields of UK High Educational institutions, the JISC-SemTech Project¹¹⁸ (carried out by the University of Southampton) developed several ontology-based applications (connected through LD) for semantic search.

Within the MIT Libraries Cataloging OASIS project it was implemented the openly available Utility Tool¹¹⁹ converting MARC (Machine-Readable Cataloging Standards)¹²⁰ and MODS (Metadata Object Description Schema)¹²¹ into RDF. RDF obtained data can serve as output for a subsequent modeling of common ontology based on RDF. Also The ‘Bibliographic Ontology Specification’ (D’Arcus et al., 2009) has provided the general concepts and properties useful to publish citations and bibliographical references (e.g. books, articles) through the Semantic Web ontologies, relying on CCL licenses and RDF technologies.

112 <<http://umbel.zitgist.com/>>; <<http://virtuoso.openlinksw.com/>>; <<http://latc-project.eu/>>

113 Europeana Libraries: Aggregating digital content from Europe’s libraries, <http://ec.europa.eu/information_society/apps/projects/factsheet/index.cfm?project_ref=270933>; Europeana Libraries Proposal, <<http://www.europeana-libraries.eu/documents/868553/bc8a98bc-5339-4117-bb2b-4cb5f75b7aaf>>

114 <<http://books.google.com/>>

115 <<http://www.dart-europe.eu/basic-search.php>>

116 <<http://www.doaj.org/>>

117 University of Ghent, <<http://biblio.ugent.be/input>>; DRIVER Technology Watch report, 7th Framework Programme, <<https://biblio.ugent.be/input/download?func=downloadFile&recordId=723558&fileId=723577>>

118 <<http://www.jisc.ac.uk/media/documents/projects/semtech-report.pdf>>

119 MIT Libraries Cataloging OASIS, <<http://libstaff.mit.edu/colserv/cat/>>; Utility Tool, <<http://simile.mit.edu/repository/RDFizers/marcmods2rdf/>>

120 Metadata Object Description Schema (MODS), <<http://www.loc.gov/standards/mods/>>; Machine-Readable Cataloging Standards (MARC), <<http://www.loc.gov/marc/>>

121 <<http://www.loc.gov/standards/mods/>>

The 'Bibliographic Ontology' (Bibo)¹²² provides a good practical example of publishing bibliographic data in RDF. The Bibo may convert in RDF a wide range of current metadata formats. Another experience is an Open catalogue of the world's cultural works (books, music, films) called 'Bibliografica'¹²³, which runs on the OpenBiblio software¹²⁴. 'Bibliografica' offers an ontology-driven platform based on the native RDF linked data support, FRBR-like domain model, and Wiki-like recording of every change of bibliographic data. 'Bibliografica' allows to create personalized collections, to add additional information to bibliographic entries, and to share these last with Wikipedia.

In this context, it is also worth mentioning the 'MarcOnt'¹²⁵, which through the integrated RDF Translator provides a technique for common ontology integration of bibliographic descriptive formats such as MARC21, BibTeX and Dublin Core. Meanwhile, the released MODS Ontology¹²⁶ represents a good ontology strategy for the migration of MARC metadata into MODS¹²⁷ expressing its entries in RDF and OWL. This make bibliographic data acceptable by the producers of LD.

Hereafter, within the UCSD Libraries' Digital Library Program¹²⁸, that discovered some limits of the DSpace¹²⁹ and Fedora¹³⁰ software regarding the acceptance of some bibliographic data formats, there was developed the ARK [Archival Resource Key] tool (Kunze, 2003). This tool allows to transform hundreds thousands of MARC and MODS data in RDF, further loading it into the AllegroGraph RDF¹³¹ queried by the SPARQL language. Beyond this experience, new versions of Eprints¹³² and DSpace (Bosman, 2009) software were released allowing to publish bibliographic data in RDF formats, to customize and qualify this data through semantic Dublin Core Application Profiles such as SWAP, IAP and TMAP, and to connect different data through LD mechanisms.

122 <<http://bibliontology.com/>>

123 <<http://bibliografica.org>>

124 <<http://openbiblio.net/p/openbiblio-software/>>

125 <<http://www.marcont.org/>>

126 <<http://www.chrisfrymann.com/2009/05/21/mods-ontology/>>

127 There are different tools making different data formats available in MODS (whose current version is 3.4): MarcEdit (<<http://people.oregonstate.edu/~reaset/marcedit/html/index.php>>); using xslt from DC to MARCXML (<<http://www.loc.gov/standards/marcxml/xslt/DC2MARC21slim.xsl>>) and then use the stylesheet from MARCXML to MODS (<<http://www.loc.gov/standards/mods/v3/MARC21slim2MODS3-4.xsl>>). A (Simple) DC-to-MODS stylesheet (<<http://www.loc.gov/standards/mods/simpleDC2MODS.xsl>>). From FGDC to MODS (<<http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1174&context=libraryscience>>). The Library of Congress has a DC to MODS xsl available (<<http://www.loc.gov/standards/mods/simpleDC2MODS.xsl>>). See also LOC conversions for MODS (<<http://www.loc.gov/standards/mods/mods-conversions.html>>).

128 UCSD Libraries' Digital Library Program, <<http://libraries.ucsd.edu/about/digital-library/index.html>>

129 <<http://www.dspace.org/>>

130 <<http://www.fedora-commons.org/>>

131 AllegroGraph® RDFStore 4.2.1, <<http://www.franz.com/agraph/allegrograph/>>, <http://www.chrisfrymann.com/image/mods/rdf_graph.png>; Graph Visualization Tool (RDF Gravity). User Documentation. Knowledge Information Systems Group, Austria, <<http://semweb.salzburgresearch.at/apps/rdf-gravity/>>

132 EPrints 3.2.0 <<http://files.eprints.org/494/>>; <<http://blogs.ukoln.ac.uk/ap-support/tag/bibo/>>

In this new, and certainly not exhaustive, scenario of semantic exposure of bibliographic data on the web, it is important that more and more information providers (included cultural, scientific and administrative bodies) make their data available in formats adaptable to the Semantic Web, replicating current experiences and proposing new projects, tools and use cases.

LODe-BD: Enabling Bibliographic Data to become Linked Open Data

The concept 'LODe' refers itself to the concept of 'Bibliographic Data' (BD), forming together a complex concept such as 'LODe-BD'. Some authors¹³³ recognize for the 'LODe', particularly, for the final 'e' something that can be 'e'-mbedded within the system itself. On the other hand, the AIMS (Agricultural Information Management Standards)¹³⁴ (Subirats, Nicolai and Waltham, 2010), team defines 'LODe' as LOD-'e'-nabled, where "enabled" is the potentiality of data to become Open and Linked (LOD) data.

To publish BD as LOD data, there must be identified standards, formats and licenses able to support BD within the LOD cloud space. There should be a common agreement on the data exposure as well as "minimal set of properties meaningful in data sharing" (Subirats, and Zeng, 2011) in the LOD data space. To assist this task, the AIMS team has developed and posted on its website the Recommendations LODe-BD¹³⁵ (Figure 5). These Recommendations provide the necessary steps and assessment tools to support agents in choosing strategies and standards for encoding BD as LOD.

The screenshot shows a web page with a navigation menu on the left and a main content area on the right. The navigation menu includes links for 'FAO Home', 'AIMS Home', 'LODe-BD 1.1', '1. About', '2. Background', '3. Terminology', '4. General Recommendations', '4.1. Entities', '4.2. Properties', and '4.3. Standards'. The main content area is titled '4. General Recommendations: Entities, Properties, Standards, and Crosswalk' and contains the following text:

LODe-BD Recommendations 1.1 > 4. General Recommendations: Entities, Properties, Standards, and Crosswalk

Once a content provider has decided to publish a bibliographic database as Linked Open Data, there are important components that should be considered:

- **What kinds of entities and relationships are involved in bibliographic resource descriptions?**
The definition of a conceptual model helps to establish an overall picture of involving entities and relationships in bibliographic descriptions. In a broader context, the use of a similar conceptual model among content providers should also help to foster a common understanding of the involving data models. LODe-BD proposes a simple conceptual model based on three entities: resource, agent and theme. The model should provide sufficient capabilities for the data providers to present their content (such as document repositories and library catalogues) for sharing in the traditional environment or transferring to the Linked Data environment.
More at 4.1. Core entities and relationships for bibliographic resource description
- **What properties should be considered for publishing meaningful/useful LOD-ready bibliographic data?**
In the Linked Data context any content provider can expose anything contained in its local database. However, in the case of bibliographic data, standardized types of information should be considered in order to maximize the impact of sharing and connecting of the data. LODe-BD has identified nine groups of common properties for describing bibliographic resources. They form the backbone of LODe-BD Recommendations and are the basis of Chapter 5. About two dozen properties used for describing a bibliographic resource as well as an additional two sets of properties for describing relations between bibliographic resources or between agents are included with specific best practice recommendations.

Figure 5. A Fragment of the LODe-BD Recommendations

Source: <<http://aims.fao.org/lode/bd/entities>>

133 De Robbio, A., Giacomazzi, S. Dati aperti con LODe, «Bibliotime», XIV, 2 (July 2011), <<http://didattica.spbo.unibo.it/bibliotime/num-xiv-2/derobbio.htm#nota91>>

134 AIMS, About LODe-BD Recommendations 1.1, <<http://aims.fao.org/lode/bd/about>>

135 LODe-BD Recommendations v.1.1, <<http://aims.fao.org/lode/bd>>

Particularly, the Recommendations, based on five key principles, provide a set of instructions and tips enabling structured bibliographic data describing digital resources (such as articles, monographs, thesis, conference papers, presentation material, research reports, learning objects)¹³⁶ to acquire LOD characteristics. The five key LODE-BD principles call:

1. To promote the use of well-established metadata standards as well as the emerging LOD-enabled vocabularies proposed in the Linked Data community;
2. To encourage the use of authority data, controlled vocabularies, and syntax encoding standards whenever possible in order to enhance the quality of the interoperability and effectiveness of information exchange;
3. To encourage the use of resource URIs as names for things for data values when they are available;
4. To facilitate the decision-making process regarding data encoding for the purpose of exchange and reuse;
5. To provide a reference support that is open for suggestions of new properties and metadata terms according to the needs of the Linked Data community.

The LODE-BD Recommendations not only provide information on how to publish and use open bibliographic data as LD, but also on where to retrieve LD sets and vocabularies supporting the LD publishing¹³⁷.

¹³⁶ The Recommendations LODE-BD may be extended to accommodate other kinds of information resources in the future. LODE-BD is a part of a series of LODE recommendations overarching a wide range of resource types including, in addition to this document for bibliographic data, the encoding of value vocabularies used in describing agents, places, and topics in bibliographic data.

¹³⁷ LODE-BD Recommendations. Step Forward. References and Links (How to publish and consume Linked Data, <<http://aims.fao.org/lode/bd/step-forward>>)

As a first step enabling BD to move towards LOD, the Recommendations offer the descriptive guide about necessary properties of bibliographic metadata, arranging them in nine groups (Table I).

Table I. Groups of Common Properties (LODe-BD) for Bibliographic Data

1.	Metadata	Property/description
	Title	Title is one of the most important and relevant access to any resource. The information is usually provided through a series of properties, including the title, alternate title, subtitles, translated titles.
2.	Responsible entity (Creator, Contributor, Editor)	This group contains the properties associated with the Agent (Creator, Contributor, Publisher) responsible for creating and publishing of the resource content.
3.	Physical characteristics	Properties that describe the appearance and characteristics of the physical form of a resource. They are: date, identifier, language, format, edition / version.
4.	Collocation	It is considered important for a resource to be located and retrieved in the area of information exchange. The properties of this group are represented by location and availability of a resource.
5.	Subject	In contrast to the physical characteristics, the group «Subject» encompasses properties describing or, at least, helping to identify what a resource denotes (under the subject term, classes/categories, keywords assigned geographic entity).
6.	Description of the Content	There are two main types of descriptions being focused on the content of a resource, rather than on the physical object: a) every description is representative of the content, usually in the form of abstract, summary, notes, table of contents; b) the type or kind of resource.
7.	Intellectual property	Any property that is concerned with the intellectual property rights relating to access and use of a resource, with particular regard to the rights, use and access conditions.
8.	Use	The properties that relate to the use of a resource, rather than the characteristics of the resource itself. Typical characteristics are: users, their level of education.
9.	Relationship between documents/ agents (responsible for the creation / publication of documents)	This group defines the relations/connections between two resources or between two agents. Considering the significant number of properties of the connections, the specific properties of the relations are explained in other parts of the Recommendations.

Source: <<http://aims.fao.org/lode/bd/properties>>

The nine groups are further extended with the specific properties, presented and explained in Table II.

Table II. Groups of Properties Describing Bibliographic Resources

A Groups of properties	B Property	C Requirement M HR R O		D Recommendation on the control of values Value Control	E important attributes associated with individual properties
		Non Analytical	Analytical		
		bibliographic descriptions			
1. Title Information	title++ mandatory property	M	M	n (not controlled should use a name authority or a controlled vocabulary or should follow a syntax encoding rule.	language
	alternative title	O	O	n	
	title supplement	O	O	n	
2. Responsible Body	creator+	HR	HR	n or Name authority (personal, corporate body, conference)	
	contributor	O	O	n or Name authority	scheme can be either a value encoding scheme or a syntax encoding scheme.
	publisher/issuer+ highly recommended property	HR	R	n or Name authority	
3. Physical Characteristics	date++	M	M	Syntax encoding rule	scheme
	identifier+	HR	HR	Syntax encoding rule	scheme
	language++	M	M	Controlled list	scheme
	format/medium+	HR	HR	Controlled list	scheme
	edition / version	R	R	n	
	source+	HR	R	n	
4. Location	location++	M	M	n or Rule [Holding unit names may be managed through a controlled list]	
	availability	O	O	n	
5. Subject	subject term+	HR	HR	Controlled vocabulary	language scheme
	classification	O	O	Controlled vocabulary, Classification system	scheme
	[freely assigned] keyword	R	R	n	language
6. Description of content	geographic term	O	O	Controlled vocabulary	language scheme
	description/abstract (or note/summary/ table of contents)	R	R	n	language
	type/form/genre	R	R	Controlled vocabulary	language scheme
7. Intellectual property	rights+ term of use access condition	R	R	n [Rights holders may be managed through name authorities]	
8. Usage	audience	O	O	Controlled list	scheme
	literary indication	O	O	Controlled list	scheme
	education level	O	O	Controlled list	scheme
9. Relation	[relation between resources]+	O	HR	Controlled resource IDs	
	[relation between agents]	O	O	n or Name authority	

LODe-BD Decision Trees

The decision trees of LODe-BD are designed to assist any bibliographic data provider in the metadata selection process. Particularly, these decision trees are based on flow charts that guide the choice of properties included in the already mentioned nine groups of metadata. Moving from the property describing a resource instance, each diagram shows the flowchart for a decision point, offering a progressive solution for encoding metadata (by means of symbols presented together with their description in Figure 6).

Name	Symbol	Used in Flowchart
Narrow Oval		Beginning of a decision tree
Flow-line		Direction of logic flow
Dashed Flow-line		Suggested direction of logic flow
Diamond		A decision to be made
Rectangle		A process to be carried out
Parallelogram		An input or available info, sources
Oval		End of a decision

Figure 6. Symbols and Their Definitions in the Flowcharts LODe-BD

Once having the tools and their descriptions in the Figures and Tables presented, it is important to provide passages useful for data developers to understand how to manage the available data. These passages are graphically explicated through the Concept Maps and flowcharts. In Figure 7 the decision trees and the explanation how to encode the properties of "Title" and "Creator" are reported.

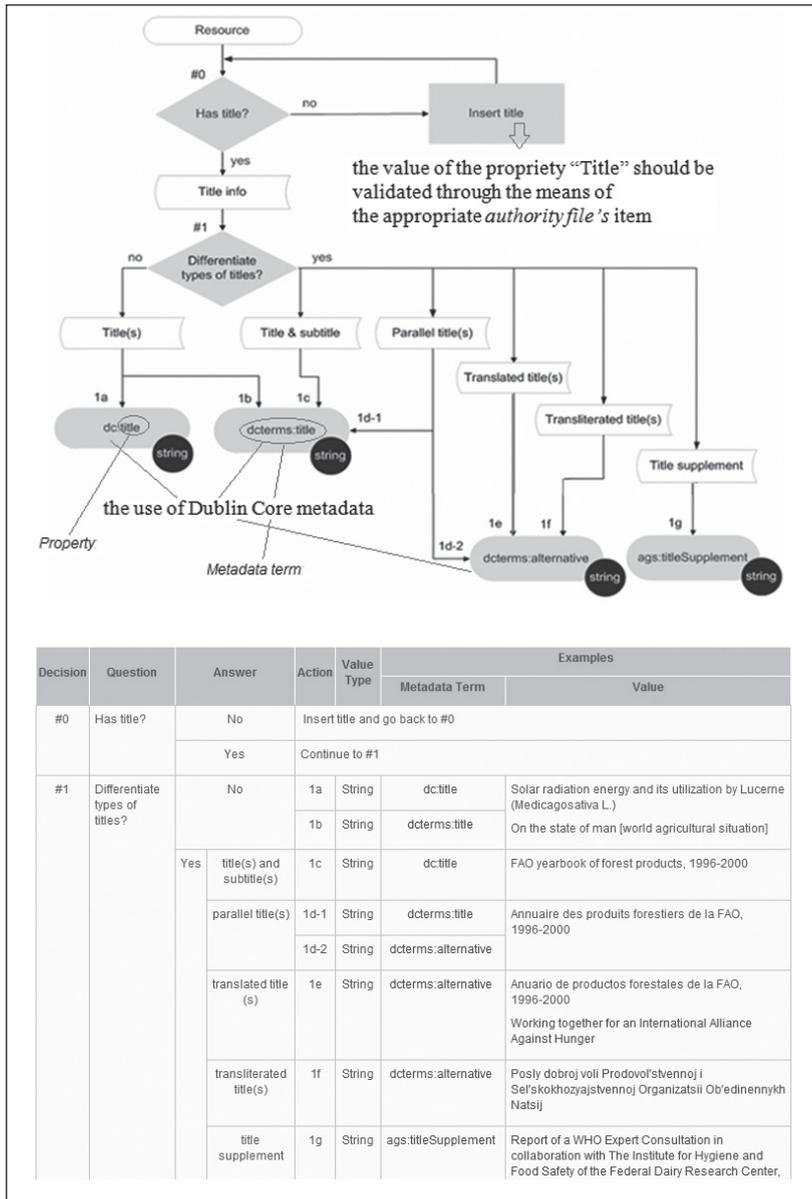


Figure 7. Encoding the Properties of "Title" and "Creator"

Source: <<http://aims.fao.org/lode/bd/title>>; <<http://aims.fao.org/lode/bd/creator>>

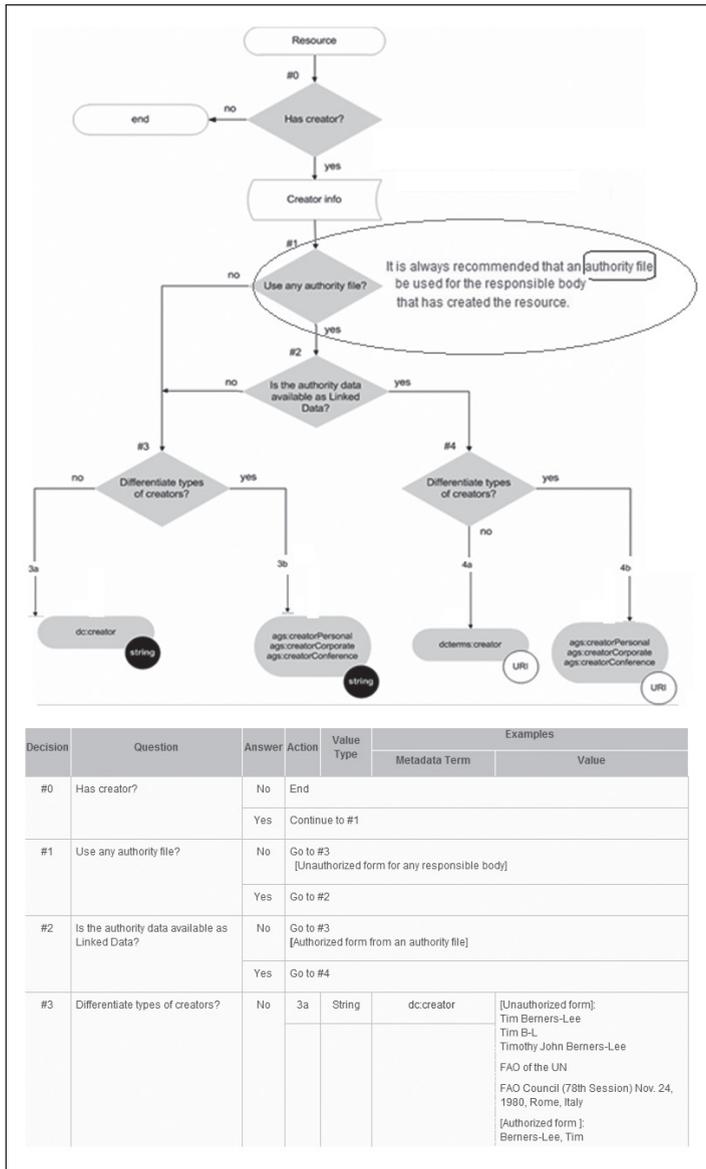


Figure 7. Encoding the Properties of "Title" and "Creator"

Source: <<http://aims.fao.org/lode/bd/title>>; <<http://aims.fao.org/lode/bd/creator>>

Each decision tree provided for each property and explained in its details is designed to facilitate the selection of the appropriate strategies to design semantic data models validated by means of authority control and based on the standards proper to the communities involved in bibliographic data management. Moreover, such a design should take into consideration the concept of total openness (Open Data) and accessibility of LOD-e BD.

Once it is decided to publish a bibliographic database in the LOD modality, there should be decided what types of entities and relations are involved in the description of bibliographic resources. For this purpose, a LODE-BD concept model is introduced (Figure 8). It aims at sharing common understanding on creating entities and relations of semantically rich BD. The LODE-BD concept model is developed on the basis of the FRBR (Functional Requirements for Bibliographic Records) (Saur, 1998), by means of which it is possible to extend and reconsider significantly the LODE-BD strategies in modeling semantic metadata.

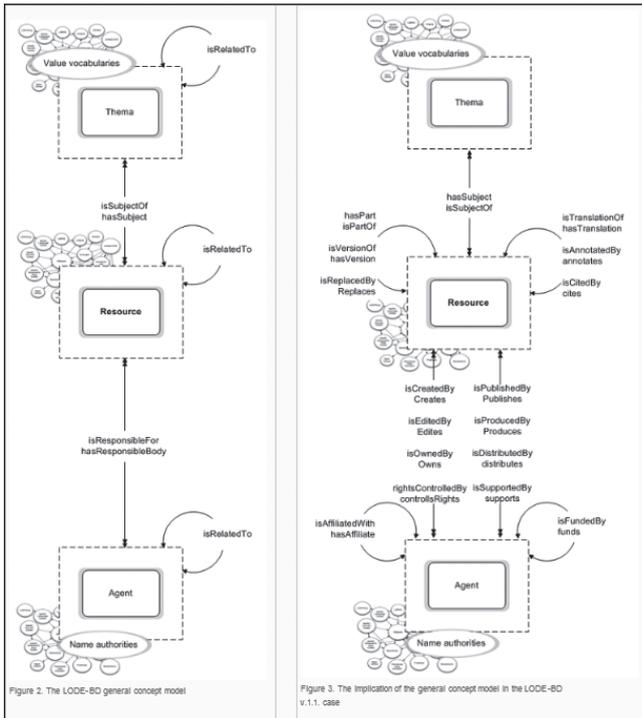


Figure 8. LODE-BD Concept Model: Defining the Subject, Theme, Agents and their Relationships

Source: Subirats, and Zeng, 2011.

In the left part of Figure 8 there is provided a high level abstraction of LODE-BD general concept model describing the central entity of any information system – Resource, related to Theme and Agent. In the right part of Figure 8, the implication of the general concept model in the LODE-BD is shown and the examples of possible relationships among instances of different entities are provided:

1. The entity Resource is the starting point of any bibliographic description in LODE-BD decision trees.
2. Relationships are established among the entity Resource and other two major entities: Agent: the entity responsible for creating of the content and/or for dissemination of the resource; and Theme (subjects, themes/topics, concepts and categories of the created content).
3. There may be also created relationships among instances of an Entity. For example, a Resource can be connected to another Resource. An Agent may be related to another Agent.
4. The relations between any pair of instances vary and may be created at different levels. For example, an Agent can provide funds for the creation of an original work, for the translation of this work, or for the release of a new format of the translation.
5. Authority Control (name authorities, value vocabularies) is considered an important element of the model. Agents, regardless of their role in the relation to a Resource, should be managed through authority files of names. In the same manner, through appropriate value vocabularies, title, main concepts (themes/topics), and geographical locations of the Resource should be controlled. Different authority files are already available in the LOD cloud.

The LODE-BD concept model represents one of the best practices enabling Bibliographic Data to get ready as LOD data. This model can also be used to mark internal, external and collaborative responsibilities of a LOD-enabled project highlighting each of its phases.

Standards for Metadata LOD-Ready

The Recommendations LODE-BD list widely-used metadata standards and emerging LOD-enabled vocabularies, that should be used to set up high-quality “LOD-ready metadata” (Table III)¹³⁸. Despite the fact that the standards selected in the LODE-BD are focused on the knowledge domain supporting the Agriculture sector (AGMES), any other community modeling its knowledge datasets can adopt the LODE-BD as a reference model. It is only a matter of fact to select another list of standards appropriately.

138 Standards for publishing vocabularies/ontologies in Linked Data (e.g. SKOS) are not included in the table below and will be included in other recommendations LODE.

Table III. Metadata Standards and Emerging LOD-Enabled Vocabularies

dc:	Dublin Core Metadata Element Set Dublin Core Metadata Element Set (DCMES or DC) Dublin Core Metadata Initiative (DCMI) Namespace: http://purl.org/dc/elements/1.1/ Page: http://dublincore.org/documents/does/ = ISO 15838 The Dublin Core Metadata Element Set is a vocabulary of fifteen properties for use in resource description. The fifteen Dublin Core elements in this stan are part of a larger set of metadata vocabularies and technical specifications maintained by the Dublin Core Metadata Initiative (DCMI).
dcterms:	DCMI Metadata Terms DCMI Metadata Terms Dublin Core Metadata Initiative (DCMI) Namespace: http://purl.org/dc/terms/ Page: http://dublincore.org/documents/dcmi-terms/ The DCMI Metadata Terms is an authoritative specification of all metadata terms maintained by DCMI. As a full set of DCMI vocabularies it also include sets of resource classes (including the DCMI Type Vocabulary), vocabulary encoding schemes, and syntax encoding schemes.
bibo:	Bibliographic Ontology Bibliographic Ontology Specification Group Namespace: http://purl.org/ontology/bibo/ Page: http://bibliontology.com/specification The Bibliographic Ontology is designed for use in describing bibliographic things on the semantic Web in RDF. One of the usages of the ontology is to de any kind of document in RDF, in addition to other usages such as being a citation ontology or as a document classification ontology.
eprint:	Eprints Terms, UKOLN Namespace: http://purl.org/eprint/terms/ Page: http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Terms The Eprints Terms include eprints-specific metadata properties and encoding schemes that have been created as part of the Dublin Core-based Scholarly Works Application Profile.
marcrel:	MARC List for Relators Library of Congress Namespace: http://id.loc.gov/vocabulary/relators/ Page: http://www.loc.gov/marc/relators/relators.html Relator terms and their associated codes are originally designed for use with the MARC records, for designating the relationship between a name and a bibliographic resource.
agls:	AGLS Metadata Standard (Australian Government Locator Service) Namespace: http://www.agls.gov.au/agla/terms/ Page: http://www.agls.gov.au/documents/aglsterms/ = Australian Standard AS 5044-2010. The AGLS Metadata Standard is developed to promote consistency of discovery of government resources. It provides a set of metadata properties and associated usage guidelines to improve the visibility, manageability and interoperability of online information and services.
ags:	AgMES (Agricultural Metadata Element set) Namespace: http://purl.org/agmes/1.1/ Page: http://aims.fao.org/website/AgMES-1.1-Namespae-Specification/sub AgMES is an application profile of the Dublin Core Metadata Element Set in the domain of agriculture with respect to description, resource discovery, interoperability and data exchange for different types of information resources. (The specification is to be updated. Metadata terms used in this report are from the current version.)

Source: <<http://aims.fao.org/lode/bd/metadata-standards>>

The selection of the appropriate standards to get different metadata LOD-ready should fall on the choice of standards widely used in the reference community, as well as on the LODe vocabularies becoming increasingly popular within the same community. To guide the choice of right standards, the role of the “Decision Trees” approach remains specifically important providing assistance in the selection process presented through flowcharts and identifying the relevant properties in each of the nine groups¹³⁹ of metadata for LODe-BD.

¹³⁹ De Robbio, A., Giacomazzi, S. Dati aperti con LODe, «Bibliotime», XIV, n.2 (2011), <<http://didattica.spbo.unibo.it/bibliotime/num-xiv-2/derobbio.htm#nota91>>

Library Linked Data: From Collections to Connections

The continuous efforts of the W3C LLD XG Library Linked data incubator group¹⁴⁰ in exploration and creation of harmonized theoretical and practical issues guiding the implementation of connected bibliographical data and their collections on the web of data (Hanneman, 2010; Karen, 2010), have been resulted into development of recommendations concerning the publishing of Library legacy data as Linked Data¹⁴¹. Particularly, the "Library Linked Data Incubator Group Final Report"¹⁴², together with the supporting documents "Report on Use Cases"¹⁴³ and "Report on Datasets, Value Vocabularies, and Metadata Element Sets"¹⁴⁴, released by the LLD XG in 2011, can be seen as a 'meeting place' among documented ideas of different professional communities¹⁴⁵ on publishing open bibliographic data as linked data under CKAN MetaData Conventions¹⁴⁶.

The graphical version of various datasets connected in the space of Library Linked Data (LLD) (Figure 9) is developed by the CKAN LLD group¹⁴⁷. Although this graphical vision has still not been effectively 'translated' into a real Linked Data service, the conceptualization of different aspects of LLD represents a real effective start point to collect¹⁴⁸ and to program bibliographic data to enter into the LOD Cloud.

The colored circles of the graph represent the data packets developed within the domain LLD, while the gray circles represent the related data packets, most of which are coming from the LOD Cloud space (e.g. DBpedia, GeoNames, VIAF, DDC, LCSH, RDA, FOAF, DCMI Metadata Terms, the elements of RDA¹⁴⁹). Size of the circles and thickness of the lines reflect the size of the packets and the number of external links provided¹⁵⁰.

140 W3C LLD XG Library Linked data incubator group, <<http://www.w3.org/2005/Incubator/lld/>>

141 Wiki on LLD: <<http://www.w3.org/2001/sw/wiki/LLD>>

142 <<http://www.w3.org/2005/Incubator/lld/XGR-llid-20111025/>>

143 <<http://www.w3.org/2005/Incubator/lld/XGR-llid-usecase-20111025/>>

144 <<http://www.w3.org/2005/Incubator/lld/XGR-llid-vocabdataset-20111025/>>

145 Library Linked Data Incubator Group Charter, <<http://www.w3.org/2005/Incubator/lld/charter>>

146 <<http://www.w3.org/wiki/TaskForces/CommunityProjects/LinkedLibraryData/Datasets/CKANmetainformation>
<<http://thedatahub.org/group/llid>>

148 Guidelines for Collecting Metadata on Linked Datasets in CKAN, <<http://www.w3.org/wiki/TaskForces/CommunityProjects/LinkingOpenData/DataSets/CKANmetainformation>>

149 DBpedia, <<http://www.w3.org/2005/Incubator/lld/XGR-llid-vocabdataset-20111025/#DBpedia>>; GeoNames, <<http://www.w3.org/2005/Incubator/lld/XGR-llid-vocabdataset-20111025/#GeoNames>>; Virtual International Authority File (VIAF), <[http://www.w3.org/2005/Incubator/lld/wiki/Use_Case_Virtual_International_Authority_File_\(VIAF\)](http://www.w3.org/2005/Incubator/lld/wiki/Use_Case_Virtual_International_Authority_File_(VIAF))>; Library of Congress Subject Headings linked data, an experimental service that makes the Library of Congress Subject Headings linked data, <<http://lcsjsubjects.org/>; <http://www.w3.org/2005/Incubator/lld/XGR-llid-vocabdataset-20111025/#Published_value_vocabularies>; Metadata element sets such as DCMI Metadata Terms, the elements of RDA: Resource Description and Access, and the Friend of a Friend vocabulary (FOAF), <http://www.w3.org/2005/Incubator/lld/XGR-llid-vocabdataset-20111025/#Metadata_Element_Sets>

150 Vatan, B. Porting library vocabularies to the Semantic Web, and back. A win-win round trip. In: IFLA 2010 Libraries and the semantic web, cit., <<http://www.ifla.org/files/hq/papers/ifla76/149-vatant-en.pdf>>

Furthermore, by integrating into the Semantic Web, bibliographic datasets could interact widely with various web ontologies (Bai et al., 2011; Westrum et al., 2012) (e.g. SKOS, Library of Congress Name Authority File, BIBO, Lexvo.org ontology, CC REL, OPM, UMBEL vocabulary, WEMI FRBR ontology, ontology bibliography NSTL¹⁵⁴) drawing automatically from them the descriptions to enrich semantically their properties and connections.

Generating rich clusters containing bibliographic data and their collections, as well as their descriptive and normalizing properties, connected through the LD mechanisms, it will be of course also possible to process a variety of sophisticated queries on those clusters. Moreover, semantics of data connected by LD makes it easier to take advantage of a service that only partially matches a query request, involving the creation of a «value chain» in which sub-assemblies of information are passed from one agent to another, each one «adding value,» to construct the final product requested by the user query.

Despite the obvious benefits of LD¹⁵⁵, the numerous projects¹⁵⁶ promoting the creation of innovative semantic services (Nandzik et al., 2010), the communities aiming at publishing their data on the web as LD should follow an accurate specific and decisive coordination, in order to counter the risk of irregularly implementation of these technologies.

Among the main obstacles to publication of data with the LD mechanisms, there is still reluctance on the different organizations to adopt new software platforms, as well as the lack of staff with adequate knowledge and skills to use these platforms before and after their effective implementation.¹⁵⁷

If professionally trained librarians, as well as experts developing and managing metadata, may find themselves feeling overwhelmed by the prospect of producing semantic metadata and maintaining ontologies, other communities operating on the web may find themselves even more frustrated to support the similar tasks.

154 Existing published Vocabularies available for reuse, <http://www.w3.org/2005/Incubator/ld/wiki/Vocabularies#Existing_published_Vocabularies_available_for_reuse>; Value Vocabularies: through SKOS to Library Linked Data, <<http://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2011/>>; <http://www.w3.org/2005/Incubator/ld/XGR-ld-vocabdataset-20111025/#Library_of_Congress_Name_Authority_File_.28LC.2FNAF.29>; "In a Linked Data world without the WEMI ontology, we have a mishmash of data and no way to know what 'thing' the data is really about", <<http://bibwild.wordpress.com/2012/02/15/yet-another-defense-of-frbr-in-a-linked-data-world/>>; FRBR WEMI ontology: <<https://bibwild.wordpress.com/2012/02/21/data-modelling-and-frbr-wemi-ontology/>>

155 LD, being much more flexible than ontological schemes such as Universal Plug and Play, such a semantic approach opens up a world of exciting possibilities. LD are potential enough to generate knowledge organization systems that can be used to organize data within a larger variety of different fields and areas of interest. "What are the enterprise benefits of Linked Data?, (Why adopt it?)", AI 3, <<http://www.mkbergman.com/447/what-is-linked-data/>>

156 Net7, Semantic Web, <<http://netseven.wp.netseven.it/servizi/semantic%C2%A0web/>>

157 Jackson, J. Berners-Lee, T. Machine-readable Web still a ways off, Government Computer News, 30 Oct., 2009, <<http://gcn.com/articles/2009/10/30/berners-lee-semantic-web.aspx>>

There are also tensions between open standards¹⁵⁸, and open publishing¹⁵⁹ required for the Semantic Web and existing methods supporting and maintaining the of data management infrastructures. From a more positive perspective, different web communities could certainly imitate the already existing good practices and use cases for publishing their bibliographic data with LD technologies.

As to the conclusions of the LLD Report, they are the following:

- ◇ bibliographic data are still not integrated with web resources;
- ◇ library standards have been designed only for the library community;
- ◇ bibliographic data are expressed primarily in natural textual language;
- ◇ the library community and the Semantic Web use different terminology for similar concepts of metadata;
- ◇ technological changes concerning library depend on the development of the largest vendors of digital information systems.

To counter the difficulties of integration of bibliographic data in the Semantic Web, the Report provides four main recommendations, with which:

1. invites library communities to identify their data packets (including thesauri and controlled vocabularies) (Vatant, 2010) to be published in LD mode and through Semantic Web standard languages (e.g. SKOS) in order to prepare them for a successful integration in the Semantic Web space, as well for creating of new and potentially unlimited services. Moreover, the Report invites libraries to spread their know-how on freeing data (Coyle, 2010; Danowski, 2010)¹⁶⁰ and in implementing of widely-used licenses¹⁶¹ allowing libraries to store, share, enrich and re-use data openly as needed. The Report identifies three potential categories of bibliographic data to be published in LD and to be managed through appropriate namespace¹⁶². They are:
 - a. dataset forming the triples describing bibliographic resources, people, organizations and related entities;
 - b. controlled (value) vocabularies from which there can derive the triples formed by values, properties and attributes for authority files;
 - c. formats and vocabularies (schemas) of metadata generating the triples describing attributes and relationships within the bibliographic datasets and controlled vocabularies.

158 Open Standards Consultation, <<http://consultation.cabinetoffice.gov.uk/openstandards/>>

159 <http://en.wikipedia.org/wiki/Open_publishing>

160 Hyperborea e l'Open Data, <<http://www.hyperborea.com/web/guest/48>>

161 Because each metadata set may have individual legal and privacy characteristics, appropriate licenses are designed on an individual dataset basis. However, the goal is to make these licenses as broad as possible.

162 British National Bibliography (BNB) uses (MARC, VIAF, ISBD, DC, LCSH, Lexvo, GeoNames ecc.) namespaces, <<http://thedatahub.org/it/dataset/bluk-bnb>>

4. The report calls upon the organizations¹⁶³ producing standards for libraries. Particularly, it invites them to increase the participation of libraries in the standardization of the Semantic Web (Dunsire, 2010), and in developing standards for bibliographic data compatible with LD. Moreover, it invites the same organizations to influence vendors (commercial companies providing digital resources and online databases) to consider and adopt these standards.
5. The report is addressing to developers of digital information systems to project high-level data management services based on the capabilities of the Semantic Web and LD. It also encourages to create URIs (even experimentally) for the elements of bibliographic datasets; to develop policies to manage effectively RDF vocabularies and related URIs; to express bibliographic datasets in terms of their reuse or mapping in connection to existing vocabularies published in LD.
6. Invite the communities of librarians and archivists: to preserve their datasets and controlled vocabularies in LD elements; to apply their know-how - referring to data management and long-term preservation – to datasets published in LD, so that LD triples become immutable, unique identifiers URI remain permanent, and names, titles, subject headings - as artifacts of cultural heritage - remain stable over time.

Observing the progress of the Semantic Web technologies, regional and national libraries together with archival institutions have a good opportunity to enrich their existing traditional roles of managers of the bibliographic universe with a new quality of the competition authorities¹⁶⁴ in the long-term preservation of data sets based on LD, in relation to cultural heritage.

Conclusions

The first step towards a web, that contains bibliographic data whose semantics is interpretable by machines, is surely to adopt an open standard format such as RDF that can be used to collect and categorize information scattered in the digital environment. The RDF will provide the data in their purest form (Raw Data¹⁶⁵: The data free from formats) and will allow it to be structured and connected (through LD) in a homogeneous way within the space of the Web of Data.

A financial investment and that in know-how of experts from different professional areas, concerning digital data generation and management, will surely improve the quality of semantic technologies and architectures of LD. It will also certainly generate

163 IFLA (Namespace Technical Group reposting to Committee on Standards); JSC for Development of RDA (DCMI/RDA Task Group, Dublin Core Metadata Initiative); DCMI Bibliographic Metadata Task Group (DCMI Vocabulary Management Community)

164 "Libraries and librarians have a very high "trust" factor with the general public, and we are in a position to raise the general quality of the Semantic Web by ensuring Library data competing with data from the widest range of sources", Dunsire, cit., 2012, http://www.nb.admin.ch/aktuelles/ausstellungen_und_veranstaltungen/00726/01611/03953/03958/index.html?lang=en

165 Berners-Lee. T., Talk on TED 2009, <<http://www.ted.com/pages/about>>

a positive return - to harmonization of different formats of web-based data, to their reuse by different information systems and web communities, and to creation of new and powerful search tools for information retrieval through targeted semantic queries - over the medium to long term.

The challenge for managers of the bibliographic universe is to deepen the knowledge of LD, to know all benefits that LD could have for libraries, to promote the knowledge of LODe-BD and to advance in the Library Linked Data implementation. Currently, these themes represent a growing process aiming at making different communities more efficient and successful in aligning their data models and policies with the Web of data. The list of scenarios that could potentially benefit from the Semantic Web technologies and LD, as they continue to evolve, is limited only by the imagination. Considering the library community as a whole in a Semantic Web scenario, one authentic concern to it would be this one: "The technology is finally ready; it's critical for libraries to begin preparations to become full participants in the world of Linked Data" (Byrne et al., 2010), "the first practical expression of the Semantic Web, useful and doable today, and applicable to all forms of data".¹⁶⁶

References

- Alexander, K., Cyganiak, R., Hausenblas, M. and Zhao, J. (2009). *Describing Linked Datasets*. 2nd Workshop on Linked Data on the Web, April 20, 2009, Madrid, Spain. Retrieved from <http://events.linkedata.org/ldow2009/>
- Bai, H., Qiao, X. and Liang, B. (2011). *Semantic bibliography based on ontology and Linked Data*. *International Conference on Dublin Core and Metadata Applications*. Retrieved from <http://dcevents.dublincore.org/index.php/IntConf/dc-2011/paper/view/20>
- Baker, T. (2010). *Designing Interoperable Metadata on Linked Data principles: Tutorial: Dublin Core – Building blocks for interoperability*. 2010 Webinars: Dublin Core: The Road from Metadata Formats to Linked Data, August, 25. Retrieved from http://dublincore.org/resources/training/NISO_Webinar_20100825/dcmi-webinar-03.pdf
- Berrueta, D. and Phipps, J. (2008). *Best practice recipes for publishing RDF vocabularies*. W3C Working Group Note. Retrieved from <http://www.w3.org/TR/swbp-vocab-pub/>
- Bechhofer, S., Ainsworth, J., Bhagat, J., Buchan, I., Couch, P. and Cruickshank, D. (2010). *Why Linked Data is not enough for scientists*. 6th IEEE e-Science Conference, Brisbane, Australia. Retrieved from <http://eprints.ecs.soton.ac.uk/21587/>
- Berners-Lee, T. (1997). *Axioms of Web Architecture: Metadata*. Retrieved from <http://www.w3.org/DesignIssues/Metadata>
- Berners-Lee, T. (1998). *Semantic Web Road map*. Retrieved from <http://www.w3.org/DesignIssues/Semantic.html>

¹⁶⁶ Linked Data FAQ, <http://structureddynamics.com/linked_data.html>

- Berners-Lee Tim, Hendler J. and Lassila O. (2001). The Semantic Web. *Scientific American*, 284(5), 34-43. Retrieved from <http://www.scientificamerican.com/2001/0501issue/0501berns-lee.html>
- Berners-Lee, T., Fielding, R. and Masinter, L. (2005). *RFC 3986: Uniform Resource Identifier (URI): Generic Syntax*. Internet Engineering Task Force (IETF). Retrieved from <http://www.ietf.org/rfc/rfc3986.txt>
- Bizer, C., Heath T. and Berners-Lee, T. (2009). Linked Data - The story so far. In: Heath, T., Hepp, M., Bizer, C. (Eds.). *Special Issue on Linked Data, International Journal on Semantic Web and Information Systems (IJSWIS)*. Retrieved from <http://linkeddata.org/docs/ijswis-special-issue>
- Bizer, C., Heath, T., Idehen, K. and Berners-Lee, T. (2008). *Linked Data on the Web*. Proceeding of the 17th International Conference on World Wide Web, Beijing, China. Retrieved from <http://www2008.org/papers/pdf/p1265-bizer.pdf>
- Bizer, C., Cyganiak, R. and Heath, T. (2007). *How to publish Linked Data on the Web*. Retrieved from <http://www4.wiwiw.fu-berlin.de/bizer/pub/LinkedDataTutorial>
- Bizer, C. and Cyganiak, R. (2009). *D2R server: Publishing relational databases on the Semantic Web*. Retrieved from <http://www4.wiwiw.fu-berlin.de/bizer/pub/Bizer-Cyganiak-D2R-Server-ISWC2006.pdf>
- Bleiholder, J. Naumann, F. (2008). Data fusion. *ACM Computing Surveys*, 41(1), 1-41. Retrieved from http://www.bioinf.jku.at/teaching/ss2011/se-inf/data_fusion_a1-bleiholder.pdf
- Byrne, G. and Goddard, L. (2010). The strongest link: Libraries and Linked Data. *D-Lib Magazine*, 16(11/12). Retrieved from <http://www.dlib.org/dlib/november10/byrne/11byrne.html>
- Berners-Lee. T. (2009). *Linked Data: Talk on TED ideas worth spreading*. Retrieved from http://www.ted.com/talks/tim_berners_lee_on_the_next_web.html
- Bosman, B. (2009). *DSpace 2.0 Demonstration*. 4th International Conference on Open Repositories, May 20. Retrieved from <http://smartech.gatech.edu/handle/1853/28078>
- Buitenhuis, J., Zelenika, I. and Pearce, J.M. (2010). *Open design-based strategies to enhance appropriate technology development*. Proceedings of the 14th Annual National Collegiate Inventors and Innovators Alliance Conference: Open, March 25-27th, San Francisco, USA. Retrieved from <http://nciaa.org/sites/default/files/pearce.pdf>
- Campbell, L.M. and MacNeill, S. (2010). *The Semantic Web, Linked and Open Data: A Briefing Paper*. JISC Cetus, Centre for Education Technology & Interoperability Standards. Retrieved from http://wiki.cetus.ac.uk/images/1/1a/The_Semantic_Web.pdf
- Corlosquet, S. and Clark, L. (2010). *The story of RDF in Drupal 7 and what it means for the Web at large*. DrupalConSF, April, 19-21, San Francisco, USA. Retrieved from <http://sf2010.drupal.org/conference/sessions/story-rdf-drupal7-and-what-it-means-web-large>
- Cyganiak, R. and Bizer, C. (2008). *Pubby: A Linked Data frontend for SPARQL endpoints*. Retrieved from <http://www4.wiwiw.fu-berlin.de/pubby/>
- Cyganiak, R., Delbru, R., Stenzhorn, H., Tummarello, G. and Decker, S. (2008). *Semantic sitemaps: Efficient and flexible access to datasets on the Semantic Web*. Proceedings of the 5th European Semantic Web Conference, ESWC. Retrieved from <http://richard.cyganiak.de/2008/papers/sitemaps-eswc2008.pdf>

- Danowski, P. (2010). *Step one: blow up the silo! - Open bibliographic data, the first step towards Linked Open Data*. Paper presented at the World Library and Information Congress 76th IFLA General Conference and Assembly, August 10-15, Gothenburg, Sweden. Retrieved from <http://www.ifla.org/files/hq/papers/ifla76/149-danowski-en.pdf>
- D'Aquin, M., Motta, E., Sabou, M., Angeletou, S., Gridinoc, L., Lopez, V. and Guidi, D. (2008). Towards a new generation of Semantic Web Applications. *IEEE Intelligent Systems*, 23(3), 20–28. Retrieved from <http://oro.open.ac.uk/23498/>
- D'Arcus, B. and Giasson, F. (2009). *Bibliographic ontology specification*. Retrieved from <http://bibliontology.com/specification>
- Doerr, M., Gradmann, S., Henniecke, S., Isaac, A., Meghini, C., and Sompel, H. (2010). *Europeana Data Model (EDM)*. Paper presented at the World Library and Information Congress 76th IFLA General Conference and Assembly, August 10-15, Gothenburg, Sweden. Retrieved from <http://www.ifla.org/files/hq/papers/ifla76/149-doerr-en.pdf>
- Dunsire, G. and Willer, M. (2010). *Initiatives to make standard library metadata models and structures available to the Semantic Web*. Paper presented at the World Library and Information Congress 76th IFLA General Conference and Assembly, August 10-15, Gothenburg, Sweden. Retrieved from <http://www.ifla.org/files/hq/papers/ifla76/149-dunsire-en.pdf>
- Dunsire, G. (2012). Linked Data, libraries and the Semantic Web. *Library science talk*, March 12-13, Geneva-Bern. Retrieved from <http://www.gordondunsire.com/pubs/pres/linkeddatalibsemweb.ppt>
- Goddard L. and Byrne G. (2010). Linked data tools: Semantic Web for the masses. *First Monday*, 15(11). Retrieved from <http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/3120/2633#2>
- Gruber, Thomas R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199–220. Retrieved from <http://tomgruber.org/writing/ontolingua-kaj-1993.pdf>
- Hanneman, J. and Kett, J. (2010). *Linked data for libraries*. Paper presented at the World Library and Information Congress 76th IFLA General Conference and Assembly, August 10-15, Gothenburg, Sweden. Retrieved from <http://www.ifla.org/files/hq/papers/ifla76/149-hannemann-en.pdf>
- Hassanzadeh, O. (2009). *A Declarative framework for semantic link discovery over relational data*. Poster at 18th World Wide Web Conference, WWW2009, April 20-24, Madrid, Spain. Retrieved from http://www.hpl.hp.com/people/min_wang6/papers/WWW_09_poster.pdf
- Heath, T. and Motta, E. (2008). Revyu: Linking reviews and ratings into the Web of Data. *Web Semantics: Science, Services and Agents on the World Wide Web* 6, 266–273. Retrieved from ScienceDirect website.
- Heath, T. and Bizer, C. (2011). *Linked Data: Evolving the Web into a Global Data Space*. Morgan & Claypool. Retrieved from Google Books website: <http://books.google.com.tr>
- Henneken, E.A. and Accomazzi, A. (2011). *A. Linking to Data - Effect on citation rates in astronomy*. Arxiv:11111618v1[cs.DL]. Retrieved from <http://arxiv.org/abs/1111.3618v1>

- Huynh, D., Mazzocchi, S. and Karger, D. (2005). Piggy Bank: Experience the semantic web inside your web browser. *Notes in Computer Science*, 3729, Retrieved from <http://hdl.handle.net/1721.1/29466>
- Karen, C. (2010a). Library data in a modern context. *Library Technology Reports*, 46(1),5-13. Retrieved from <http://alatechsource.metapress.com/content/p3022442071g7655>
- Karen, C. (2010b). Changing the nature of library data. *Library Technology Reports*, 46(1).14-29. Retrieved from <http://alatechsource.metapress.com/content/p7w420j45503jk0w/fulltext.pdf>
- Keßler, M. (2010). *Linked Open Data of the German National Library*. ECO4r Workshop / LOD of DNB. Deutsche National Bibliothek. Retrieved from http://www.eco4r.org/workshop2010/eco4r_workshop2010_mirjam_kessler.pdf
- Kifer, M. (2008). Rule interchange format: The framework. *Lecture Notes in Computer Science*, 5341, 1-11. Retrieved from <http://www.springerlink.com/content/e7v2802743688216/?MUD=MP>
- Kim, P., Ng, C.K. and Lim, G. (2010). When cloud computing meets with Semantic Web: A new design for e-portfolio systems in the social media era. *British Journal of Educational Technology*, 41(6), 1018-1020. Retrieved from http://eport.uaa.alaska.edu/BJET_PRPL_Eportfolio.pdf
- Kunze, J.A. (2003). *Towards electronic persistence using ARK Identifiers*. 3rd ECDL Workshop on Web Archives, August 21, Trondheim, Norway. Retrieved from <http://bibnum.bnf.fr/ecdl/2003/proceedings.php?f=kunze>
- Lacy, L.W. (2005). *Owl: Representing information using the Web Ontology Language*. Trafford Publishing.
- Maedche, A. (2002). *Ontology learning for the semantic Web*. Kluwer Academic Publishers.
- Miller, P., Nowack, B., Boutin, G. and Tague, T. (2009). October 2009: *The Semantic Web Gang discuss RDF*. 2009, Retrieved from <http://www.developerfusion.com/media/68002/the-semantic-web-gang-discuss-rdf/>
- Mitchell, E. (2010). Using cloud services for library IT infrastructure. *Cod4Lib Journal*, (9). Retrieved from <http://journal.code4lib.org/articles/2510>
- Nandzik, J., Heß, A., Hannemann, J., Flores-Herr, N. and Bossert, K. (2010). *CONTENTUS – Towards semantic multimedia libraries*. Paper presented at the World Library and Information Congress 76th IFLA General Conference and Assembly, August 10-15, Gothenburg, Sweden. Retrieved from <http://www.ifla.org/files/hq/papers/ifla76/149-nandzik-en.pdf>
- O'Connor, M.J., Shankar, R., Nyulas, C., Tu, S. and Das, A. (2008). *Developing a Web-Based application using OWL and SWRL*. AAAI Spring Symposium 2008, March 26–28, Stanford University, California. Retrieved from <http://www.aaai.org/Papers/Symposia/Spring/2008/SS-08-01/SS08-01-012.pdf>
- Oomen, J. and Baltussen, L.B. (2012). *Sharing cultural heritage the linked open data way: Why you should sign up*. Museums and the Web 2012, April 11-14, San Diego, USA. Retrieved from http://www.museumsandtheweb.com/mw2012/papers/sharing_cultural_heritage_the_linked_open_data

- Rusu, D. Fortuna, B. and Mladenici, D. (2011). *Automatically annotating text with Linked Open Data*. 4th Linked Data on the Web (LDOW 2011), 20th World Wide Web Conference (WWW 2011), Hyderabad, India. Retrieved from <http://events.linkedata.org/ldow2011/papers/ldow2011-paper09-rusu.pdf>
- Safari, M. (2004). Metadata and the Web. *Webology*, 1(2). Retrieved from <http://www.webology.org/2004/v1n2/a7.html>
- Sanfilippo, A., Bernardi, A., Elst, L., Fernández, L.S. and Sintek, M. (2003). *Integrating ontologies for Semantic Web applications*. Enabler/Elsnet Workshop on International Roadmap for Language Resources, September, 28-29, Paris, France. Retrieved from http://www.dfki.uni-kl.de/~bernardi/News/publication-download/Integrating_Ontologies_for_Semantic_Web_applications.pdf
- Saur, K.G. (2011). *Functional Requirements for Bibliographic Records (FRBR)*. Retrieved from <http://www.ifla.org/publications/functional-requirements-for-bibliographic-records>
- Subirats, I., Nicolai, C. and Waltham, N. (2010). *FRBR/FAO Model implementation and benefits*. ASIS&T 2010 Annual Meeting, October 26. Retrieved from <http://www.slideshare.net/faoaims/frbrfao-model-implementation-and-benefits>
- Subirats, I. and Zeng, M. (2011). *Metadata: Concentrating on the data, not on the scheme*. Presentation hold in the euroCRIS Meeting, April 26-27, Bologna, Italy. Retrieved from http://www.eurocris.org/Uploads/Web20pages/members_meetings/20110520-Bologna,Italy/Italiansession-ImmaSubirats.ppt
- Summers, E., Isaac, A., Redding, C. and Krech, D. (2008). *LCSH, SKOS and Linked Data*. International Conference on Dublin Core and Metadata Applications, September 22-26, Berlin, Germany. Retrieved from <http://dcpapers.dublincore.org/ojs/pubs/article/viewFile/916/912>
- Sure, Y., Bloehdorn, S., Haase, P., Hartmann, J. and Oberle, D. (2005). *The SWRC Ontology - Semantic Web for research communities*. 12th Portuguese Conference on Artificial Intelligence, December 5-8, Covilha, Portugal. Retrieved from http://lambda.csail.mit.edu/~chet/papers/others/o/oberle/2005_swrc_baosw.pdf
- Tolk, A., Diallo, S.Y. and Turnitsa, C.D. (2007). Applying the levels of conceptual interoperability model in support of integratability, interoperability, and composability for system-of-systems engineering. *Systemics, Cybernetics and Informatics*, 5(5), 65-74. Retrieved from [http://www.iiisci.org/journal/CV\\$/sci/pdfs/P468106.pdf](http://www.iiisci.org/journal/CV$/sci/pdfs/P468106.pdf)
- Tudhope, D. and Koch, T. (2004). New applications of knowledge organization systems. *Journal of Digital Information*, 4(4). Retrieved from <http://journals.tdl.org/jodi/article/view/109/108>
- Valkeapää, O., Alm, A. and Hyvönen E. (2007). Efficient content creation on the Semantic Web Using metadata schemas with domain ontology services (System Description). E. Franconi, M. Kifer and W. May (Eds.). In *The Semantic Web : research and applications : 4th European Semantic Web Conference, ESWC 2007, Innsbruck, Austria, June 3-7, 2007 : Proceedings* (pp. 819-828). Berlin: Springer. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.98.8805&rep=rep1&type=pdf>

- Vatant, B. (2010). *Porting library vocabularies to the Semantic Web, and back. A win-win round trip*. Paper presented at the World Library and Information Congress 76th IFLA General Conference and Assembly, August 10-15, Gothenburg, Sweden. Retrieved from <http://www.ifla.org/files/hq/papers/ifla76/149-vatant-en.pdf>
- Vockner, B., (2011). *The role of ontology in data sharing within a multi-disciplinary context – An example from energy resources management domain*. INSPIRE Conference 2011, June 27-July 1, Edinburgh, UK. Retrieved from http://inspire.jrc.ec.europa.eu/events/conferences/inspire_2011/presentations/185.pdf
- Volz, J., Bizer, C., Gaedke, M. and Kobilarov, G. (2009). *Silk – A Link Discovery Framework for the Web of Data*. 2nd Workshop on Linked Data on the Web, April 20, 2009, Madrid, Spain. Retrieved from http://events.linkedata.org/ldow2009/papers/ldow2009_paper13.pdf
- Weitzner, D. (2007). Beyond secrecy: New privacy protection strategies for open information spaces. *IEEE Internet Computing*, 11(5), 94-96, Retrieved <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4305576>
- Westrum, A.L., Rekkavik A. and Tallerås, K. (2012). Improving the presentation of library data using FRBR and Linked Data. *Code 4.Lib Journal*, (16), Retrieved from <http://journal.code4lib.org/articles/6424>
- Woutersen-Windhouwer, S., Brandsma, R., Hogenaar, A., Hoogerwerf, M. and Doorenbosch, P. (2009). *Enhanced publications: Linking publications and research data in digital repositories*. Amsterdam, Amsterdam University Press. Retrieved from <http://dare.uva.nl/document/150723>
- Zeng, M.L. (2009). *Metadata and terminology registries: Synergies and differences*. International Conference on Dublin Core and Metadata Applications, October 12-16, Seoul, Korea. Retrieved from <http://www.slideshare.net/mzeng/metadata-and-terminology-registries>
- Zeng, M.L. (2008). *Metadata for Terminology / KOS Resources*. NKOS Workshop, September 11, Washington, DC, USA. Retrieved from <http://nkos.slis.kent.edu/2008workshop/MarciaZeng.pdf>