

Resources

Chapter 6 lists some basic resources on linked data standards and usage, tutorials, and readings for beginners and more advanced linked data users. The chapter also describes discussion groups and activities where readers can interact with the library linked data community.

One of the biggest challenges is keeping up with this fast-moving area. Not only are new metadata schemas and datasets appearing on a weekly basis, the W3C has about a dozen standards in progress in its Semantic Web section. The resources in this section provide background material as well as sites and activities to follow.

Standards

The primary standards for the Semantic Web are developed by the W3C, although other groups, such as the Dublin Core Metadata Initiative, are developing related standards.

World Wide Web Consortium (W3C)

The standards for the Semantic Web are developed by the W3C and are listed on the W3C Semantic Web Publications page. This includes the standards that were discussed in this report, such as RDF and OWL, and others that were not, such as Gleaning Resource Descriptions from Dialects of Languages (GRDDL) and Rule Interchange Format (RIF). Naturally, many of the standards documents are very technical and use

specialized terminology that is difficult for many people to understand. To address this, some of the key standards have a primer that explains the concepts in the standard in more accessible language.

- Publications of the W3C Semantic Web Activity, <http://www.w3.org/2001/sw/Specs>
- RDF Primer, <http://www.w3.org/TR/rdf-primer>
- OWL 2 Web Ontology Language Primer, <http://www.w3.org/TR/owl2-primer>

Dublin Core Metadata Initiative (DCMI)

Active members of the DCMI were involved in the very early development of the Semantic Web, and this has led to Semantic Web–related standards coming out from DCMI. The primary standards were the Dublin Core Abstract Model, a practical model for metadata creation using Semantic Web concepts, and the Dublin Core Application Profile, a standard that bridges the gap between the openness of linked data and need for greater control in applications. Both of these standards are being revised to bring them up-to-date with current Semantic Web developments.

- Andy Powell et al., “DCMI Abstract Model,” June 4, 2007, <http://dublincore.org/documents/abstract-model>
- Mikael Nilsson, Thomas Baker and Pete Johnston, “Singapore Framework for Dublin Core Application Profiles,” January 14, 2008, <http://dublincore.org/documents/singapore-framework>

Websites

There are some key websites that will become familiar to you rather quickly as you work with linked data. This list is slanted toward library-related sites.

Semantic Web (W3C)

<http://www.w3.org/standards/semanticweb>

This is the homepage for the W3C Semantic Web activity. This page is an entry to all aspects of the Semantic Web work and also carries the news posts (with an RSS feed available) where standards activity and events are announced.

Linked Data

<http://linkeddata.org>

The Linked Data site is the home of the Linked Data cloud, the visualization over time of the datasets and their interrelationships. The list of datasets in the LOD cloud is at <http://thedatahub.org/group/locloud>.

Library of Congress Authorities and Vocabularies

<http://id.loc.gov>

This growing site holds the authority files and controlled vocabularies generated by the Library of Congress. This includes Library of Congress Subject Headings, Library of Congress Name Authority File, and vocabularies from the MARC 21 format and LC preservation metadata.

Open Metadata Registry

<http://metadataregistry.org>

This is the site where you can find many library data element schemas and controlled lists in Semantic Web formats, including RDA, FRBRer, ISBD, and FRAD.

The Data Hub

<http://thedatahub.org>

The Data Hub is a project of the Comprehensive Knowledge Archive Network, a site for registration of openly available data in any format. A subset of that site is the Library Linked Data group at <http://thedatahub.org/group/llid>. Here you will find links to available library datasets such as those from Europeana and the national libraries of France, Germany, and Hungary, among others. Anyone with library data to share in linked data format is encouraged to register at this site so that others can find their data.

Tutorials and Primers

Fortunately there are numerous tutorials and primer-like documents on the Web. Here are a few:

SPARQL by Example

<http://www.cambridgesemantics.com/2008/09/sparql-by-example>

This tutorial allows you to do live searching of data using SPARQL and making use of a small set of SPARQL endpoints. Once you have the idea, you can formulate your own queries by following the patterns given.

RDF Primer

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

The RDF Primer was written as part of the W3C RDF standard documentation as an entry point into the complexities of the RDF language. It uses simple examples and ample illustrations to step the reader through the basic concepts of RDF. There are sections that are still marked “@@TBD@,” that is, “to be developed,” yet this is still the best place to start for a brief introduction.

Semantic Web Primer

<http://www.linkeddatatools.com/semantic-web-basics>

This site is oriented toward developers and programmers. It has a five-part tutorial that begins with an explanation of graph data and graph databases and continues through semantic modeling, RDFS and OWL, and querying with SPARQL. If you prefer your tutorial examples to be practical and using code, this is for you.

Reading

Much of the information about the Semantic Web is naturally on the Web. For anyone wanting to master this area, though, it may help to read some of the books and articles that often provide more in-depth practical information. Books quickly become out of date, but you can update the specifics after getting some basic “book learning.”

Books

Dean Allemang and James Hendler. *Semantic Web for the Working Ontologist*. Burlington, MA: Morgan Kaufmann Publishers, 2008. ISBN: 978-0123735560.

If you are comfortable with general metadata concepts and perhaps some database management technologies, then this is a good first book on Semantic Web metadata concepts. If you are just beginning, you may wish to start with the RDF Primer (above) before delving into this book.

John Hebel et al. *Semantic Web Programming*. Indianapolis, IN: Wiley, 2009. ISBN: 978-0470418017.

Although this is in the end a book for programmers, the first half talks about Semantic Web concepts and standards without requiring any programming expertise, so you can learn about RDF and OWL and see examples of uses. If you program, then the book also gets you started with some of the current Semantic Web tools: Protégé, Jena, and Pellet.

Jeffrey T. Pollock. *Semantic Web for Dummies*. Indianapolis, IN: Wiley, 2009. ISBN: 978-0470396797.

In this case, the “dummy” is a fairly seasoned programmer with good skills in relational databases and object-oriented programming. So if your technical and programming skills are good, this book will give you a lot of practical information, relating it to things you already know. It also is aimed at the business world, where terms like enterprise are frequently used. Although that might be a bit strange for library folks, Pollock addresses issues relating to scalability, search optimization, and the types of technical skills that are essential in this new environment.

Grigoris Antoniou and Frank van Harmelen. *A Semantic Web Primer, 2nd ed.* Cambridge, MA: MIT Press, 2008. ISBN: 978-0262012423.

This book covers the Semantic Web standards in some depth, but does not cover software or programming aspects. It uses technical language from the areas of formal logic that serve as the basis for the standards, yet its brief explanations of concepts like “open world assumption” and “unique names assumption” are clear and easy to understand. The book has a textbook format, with each chapter ending in a set of questions and a list of further readings. This latter may help a reader explore concepts that are treated briefly in the text.

Articles and Reports

As interest in linked data has blossomed, the number of articles published has become very large. A search of any article database will turn up many hundreds. The ones here are those that I believe will be most helpful to librarians who are reading in this area for the first time. The first entries are particularly relevant to libraries and library data.

These two reports from 2010 give technical background on linked data from a library data perspective:

Karen Coyle. “Understanding the Semantic Web: Bibliographic Data and Metadata.” *Library Technology Reports* 46, no. 1 (January 2010).

Karen Coyle. “RDA Vocabularies for a Twenty-First Century Data Environment.” *Library Technology Reports* 46, no. 2 (February 2010).

Under the auspices of the W3C, the W3C Incubator Group met for over a year to produce a report making recommendations for library linked data:

W3C Incubator Group. *Library Linked Data Incubator Group Final Report*. October 25, 2011, W3C website, <http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025>.

Numerous articles have been written about the Semantic Web. Below are some that I have found particularly useful.

Tim Berners-Lee, James Hendler, and Ora Lassila. “The Semantic Web.” *Scientific American* 284, no. 5 (May 2001): 34–43.

This is the core document for public understanding of the Semantic Web. In this 2001 article, Berners-Lee and co-authors present a vision of future information services based on linked data and Semantic Web technologies.

Natalya F. Noy and Deborah L. McGuinness. “Ontology Development 101: A Guide to Creating Your First Ontology.” Protégé website, accessed March 8, 2012, http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html.

This guide was developed at Stanford by the same people who created the Protégé ontology management software. It walks the reader through the steps needed to develop an ontology, with good advice about common pitfalls of the process.

Karen Coyle and Thomas Baker. “Guidelines for Dublin Core Application Profiles.” DCMI website, May 18, 2009, <http://dublincore.org/documents/profile-guidelines>.

This short document provides a step-by-step approach to selecting terms for a metadata schema and developing them into an application profile that is compatible with Semantic Web standards.

Conferences and Discussion Lists

There isn’t yet a clear center for keeping up with linked library data. Along with following the work coming out of the Library of Congress and other libraries active in this area, here are some venues you might wish to plug into:

DCMI Conference and Workshop Series

<http://dublincore.org/workshops>

Every year since 1995, the year of the first Dublin Core meeting at OCLC in Dublin, Ohio, there has been a conference or workshop. While the meetings of the first decade focused on the development of the DC metadata terms, the most recent conferences have become key meeting places for those interested in promoting linked data technologies.

LODLAM Meetings

<http://lod-lam.net>

In May 2010, a conference was held in San Francisco to stimulate discussion about and work on linked open data in libraries, archives, and museums. Since then, similar meetings have been held in Washington, DC, and Wellington, New Zealand. While not yet officially a series, these meetings are energizing the discussion of a linked data future.

W3C Library Linked Data Incubator Group discussion list

<http://lists.w3.org/Archives/Public/public-ldd>

The W3C Incubator Group has completed its work and issued its final report, but the public mailing list remains an active space for the discussion of linked data and libraries. The list is international in scope, and there are participants who are involved in key linked data projects and in Semantic Web standards.

ALA ALCTS/LITA Linked Library Data Interest Group

<http://connect.ala.org/node/142470>

The newly formed interest group for library linked data meets at every ALA meeting and, with time, could become a focal point for association activities in this area. Interest groups are quite fluid in nature, and the direction of this group depends entirely on the interests of those who attend and participate.

Notes
