

Cliff Landis

# A web of meaning

## Linked open data resources on the web

Librarians have been talking about the Semantic Web for years, but with the increasing adoption of linked data standards and frameworks by major information and search providers, we are finally beginning to see it grow. Academic librarians are uniquely equipped to take advantage of this growing set of technologies to reach out to our users and to make our unique resources infinitely more useful and reusable.

Linked open data is based around the idea of creating a web of open-licensed, structured data and metadata that can be processed by computers. For example, a simple “dumb link” to a website can be enriched to add semantic meaning of not just where it’s linking to, but what it is linking to. Therefore, rather than containing only a URL address, a link could include information about who created the link, when the link was created, the title of the work that it links to, the author of that work, and when the work itself was created. This information is represented in factual “triples” of subject, predicate, and object (i.e., <http://crln.arcl.org/> - TITLE – *College & Research Libraries News*). These triples can then be linked together to create an open web of facts (i.e., *College & Research Libraries News* – PUBLISHER – Association of College and Research Libraries). And as this open web of facts grows, we can search, visualize, understand, and reuse incredibly complex collections of data in new and interesting ways.

Although not exhaustive, the following resources will help you learn about the prin-

ciples behind linked open data, identify some of the ways it is already being used, see how knowledge organizations are contributing to this emerging field, and get involved with creating the next version of the web.

### Understanding the basics

- **5 Star Open Data.** Not all published datasets are created equal. Following up on one of Tim Berners-Lee’s original ideas about linked data, this website explains the concept of the five-star “ranking” system of linked open data. The more open and structured the data, the better, but it comes with costs of time and energy. Explore this site to learn about the pros and cons of each level of linked open data. *Access:* <http://5stardata.info/>.

- **Linked Data Glossary.** As with any emerging technology, there has been an explosion of various approaches, terms, standards, and ideas associated with linked open data. To keep from being overwhelmed, keep this glossary from the W3C Government Linked Data Working Group handy. It provides short descriptions of each of the terms and links out to more in-depth descriptions and standards. *Access:* <http://www.w3.org/TR/ld-glossary/>.

---

Cliff Landis is web services librarian at Georgia State University Library, email: [clifflandis@gsu.edu](mailto:clifflandis@gsu.edu)

© 2014 Cliff Landis

- **Linked Data—The Story So Far.** For a solid historical look at linked open data, read this paper by Christian Bizer, Tom Heath, and Tim Berners-Lee, originally published in the *International Journal on Semantic Web and Information Systems* in 2009. In it, the authors lay out the basic concepts and technologies of linked open data, describe the progress and applications created up to 2009, and cover several challenges that the linked open data community continues to address today. *Access:* <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf>.

- **Linked Open Data—What is it?** Geared specifically toward memory organizations like archives, this short video from Europeana provides a simple introduction to how the technology of linked open data combines with unique resources to improve the findability and usability of those resources. *Access:* <http://vimeo.com/36752317>.

- **Tim Berners-Lee: The Next Web.** In this 2009 TED talk video, the inventor of the World Wide Web explains the concepts behind linked open data and encourages viewers to release their datasets as machine-readable linked open data. *Access:* [http://www.ted.com/talks/tim\\_berners\\_lee\\_on\\_the\\_next\\_web](http://www.ted.com/talks/tim_berners_lee_on_the_next_web).

## See it in action

- **BBC Music.** The BBC has been expanding its linked data offerings since it first started connecting its online content for its coverage of the Olympics. Particularly impressive is the BBC Music website, which aggregates data from a wide variety of sources to display a wealth of information about artists, including a feed showing which BBC shows have recently played artists' songs. *Access:* <http://www.bbc.co.uk/music/>.

- **Google Knowledge Graph.** As Google transitions its search from being an information engine to a knowledge engine, users have seen the refinement of search results that differentiate between concepts (i.e., Leonardo the painter vs. Leonardo the Ninja Turtle). This is a result of the refinements of the Knowledge Graph, Google's proprietary knowledge base. *Access:* <http://www>.

[google.com/insidesearch/features/search/knowledge.html](http://google.com/insidesearch/features/search/knowledge.html).

- **HistoryPin.** Historical photos gain new levels of meaning on this website, where users provide metadata to put photos in their correct context, place, and time. The photos are then overlaid on a global map, allowing users to browse what a place looked like over time (including overlaying photos on a street view using geospatial metadata). Archives are particularly active in sharing images from their collections on HistoryPin. *Access:* <http://www.historypin.com/>.

- **Open Street Map.** This free and open-licensed global map is a collaborative creation of volunteers (think Wikipedia meets Google Maps). Since it can be quickly edited and updated, it has been used to coordinate responses to emergencies, such as rescue efforts for natural disasters and disease outbreaks like Ebola. *Access:* <http://www.openstreetmap.org/>.



- **Tim Berners-Lee: The Year Open Data Went Worldwide.** In this short 2010 TED talk follow-up video, Berners-Lee shows how the release of linked open data on the web had an impact, including how the Open Street Map project helped save lives after the 2010 earthquake in Haiti. *Access:* [http://www.ted.com/talks/tim\\_berners\\_lee\\_the\\_year\\_open\\_data\\_went\\_worldwide](http://www.ted.com/talks/tim_berners_lee_the_year_open_data_went_worldwide).

## Large knowledge bases

- **BabelNet.** In contrast to the above knowledge bases, BabelNet functions as a multilingual linked open dictionary. By combining linked data from Wikipedia and WordNet, it provides a lexicalized semantic network, clustering meanings together into sets of multilingual synonyms (item example: <http://babelnet.org/search?word=bn:03549615n&details=1>). *Access:* <http://babelnet.org/>.

- **DBpedia.org.** One of the more famous examples of linked open data, DBpedia



was created by the Free University of Berlin, the University of Leipzig, and OpenLink

Software. It extracts structured information from Wikipedia and presents it as linked open data, including links to other linked knowledge bases (item example: [http://dbpedia.org/page/College\\_&\\_Research\\_Libraries\\_News](http://dbpedia.org/page/College_&_Research_Libraries_News)). *Access:* <http://dbpedia.org/>.

- **Freebase.** Originally created by software company Metaweb, Freebase was purchased by Google in 2010. This knowledge base has more than 2.6 billion facts, and its structured data is one of the resources used to power Google's Knowledge Graph (item example: [www.freebase.com/m/03cms3s](http://www.freebase.com/m/03cms3s)). *Access:* <http://www.freebase.com/>.

- **Library of Congress Linked Data Service.** The Library of Congress has been very active in the development of linked open data standards and has provided several datasets of authority records, available in a variety of linked open data standards. (item example: <http://id.loc.gov/authorities/names/n79058275.html>). *Access:* <http://id.loc.gov/>.

- **Wikidata.** If you want an easy-to-understand knowledge base to dabble with to get a feel for how linked open data works, do so at Wikidata. This project of the Wikimedia Foundation aims to be readable by



both machines and humans, with the goal of centralizing many of the

facts and references used on Wikipedia and its sister projects (item example: [www.wikidata.org/wiki/Q5146313](http://www.wikidata.org/wiki/Q5146313)). *Access:* <http://www.wikidata.org/>.

### Linked open data in libraries, archives, and museums

- **BIBFRAME.** The Bibliographic Framework Initiative is a new model for repre-

senting bibliographic data on the web, with the goal of eventually replacing MARC 21. Although still in its adolescence, BIBFRAME's emphasis on relationships between resources means that much of the duplication encountered in MARC will be corrected for long-term savings in costs and effort. *Access:* <http://www.loc.gov/bibframe/>.



- **Karen Coyle.** Coyle is one of the leading thinkers trying to bring linked open data to libraries. A veteran author and presenter, she creates clear connections between traditional information organization in libraries and the emerging tools and trends of linked data. Coyle's three ALA Library Technology Reports on the Semantic Web, RDA, and Linked Data are must-reads. *Access:* <http://www.kcoyle.net/>.

- **Linked Jazz.** A short video shows how this project is using linked open data to show relationships in the jazz music community, all based on archival documents and oral history transcripts. Follow this by browsing the Linked Jazz Network Visualization Tool to see the cloud of relationships and influence among the more than 8,000 jazz musicians. *Access:* <http://linkedjazz.org/>.

- **LODLAM.** This website, its associated Google Groups discussion board, and its Twitter hashtag (#LODLAM) together serve as an online gathering place for the loose network of information professionals working to bring linked open data to memory organizations. Follow along for the latest news or join in the conversation. *Access:* <http://lodlam.net/>.

- **OCLC Data Strategy and Linked Data.** OCLC provides a variety of datasets and services, having included linked open data within WorldCat records earlier this year. Additionally, OCLC provides downloads of its FAST subjects and VIAF authority files; although both tools are early in their development, they show a lot of promise for the future of linked open data. *Access:* <http://www.oclc.org/data.en.html>.

## Get started

- **Data.gov.** When you're ready to start tinkering with large data, you can grab some sample datasets from this website for the federal government's open data, organized by topic. The site also includes applications created with open government datasets, such as the United States Education Dashboard. *Access:* <http://www.data.gov/>.



- **Datahub.** This free data management platform from the Open Knowledge Foundation serves as an index of datasets, allowing users to search for datasets, register their own datasets, and get updates on datasets that they are interested in. See the Library Linked Data group for a list of 47 (and growing) datasets from the LODLAM community. *Access:* <http://datahub.io/>.

- **FreeYourMetadata.org.** Largely geared toward libraries, archives, and museums, this website provides screencasts and guides that show how to use OpenRefine and other free tools to clean, reconcile, extract, and publish metadata as linked open data. The step-by-step instructions are provided with sample datasets and reconciliation points, giving hands-on practice for the tools and concepts of linked open data. *Access:* <http://freeyourmetadata.org/>.

- **LinkedData.org.** Tom Heath runs this website that aggregates links to many resources, tools, guides, and tutorials on linked data. Although not exhaustive, this website provides a good starting place for exploring the widely distributed information available on linked open data. *Access:* <http://linkeddata.org/>.

- **LinkedData-W3C Wiki.** Similar to LinkedData.org, this page in the World Wide Web Consortium's wiki provides a good starting point for exploring the vast information available on linked data, including publications, presentations, demonstrations, and

community discussion forums. *Access:* <http://www.w3.org/wiki/LinkedData>.

- **Open Refine.** This powerful, free, open source software program helps users clean up messy datasets to prepare them for publishing to the web as linked open data. Open Refine is relatively easy to use, has a robust development community, and allows for reconciliation of datasets against standard linked data vocabularies (such as dbpedia.org, freebase.com, or LCSH). *Access:* <http://openrefine.org/>.

- **School of Data.** If you're new to working with data, the School of Data's free courses



are an excellent place to start. This collaboration between P2PU and the Open

Knowledge Foundation teaches both basic and intermediate skills in working with, cleaning, extracting, and exploring data. *Access:* <http://schoolofdata.org/>.

## Publish your linked open data

- **Best Practices for Publishing Linked Data.** This Working Group Note from the Government Linked Data Working Group serves as a great checklist for creating and publishing a linked open data dataset. Published datasets require some regular maintenance to stay functional, and this guide can help users make good choices in the design of datasets to make them more sustainable and less work in the long run. *Access:* <http://www.w3.org/TR/ld-bp/>.

- **Git (and GitHub) for Data.** Version control is vital to working with datasets; since mistakes are inevitable, version control allows us to “skip back” and fix problems (even if they happened 17 steps ago) without losing the interim changes. This blog post by Rufus Pollock describes how to use the free, open source versioning control system Git to successfully manage changes to datasets over time. *Access:* <http://blog.okfn.org/2013/07/02/git-and-github-for-data/>.

*(continues on page 505)*

/liberal-studies/first-year-program/Pages/chicago-quarter.aspx; "Chicago Quarter," DePaul University Office of New Student and Family Engagement," accessed August 8, 2014, <http://studentaffairs.depaul.edu/nsfe/chicagoquarter.html>.

3. Heather Jagman, Krystal Lewis, Brent Nunn, and Scott Walter, "Financial Literacy Across the Curriculum (and Beyond): Opportunities for Academic Libraries," *C&RL News* 75, no. 5 (2014): 254-257, accessed August 9, 2014, <http://crln.acrl.org/content/75/5/254.full>.

4. Heather Jagman, et al., "I Felt Like Such a Freshman!": Creating Library Insiders" (poster presented at the ALA Annual Conference, Las Vegas, Nevada, June 26-July 1, 2014), accessed August 12, 2014, <http://libguides.depaul.edu/ala2014aia>.

5. "About TLA," DePaul University Office for Teaching, Learning, and Assessment, accessed August 9, 2014, <http://condor.depaul.edu/tla/AboutOTLA.html>.

6. "Supplemental Instruction," DePaul University Office for Teaching, Learning, and Assessment, accessed August 9, 2014, <http://condor.depaul.edu/si/>.

7. "Assessment," DePaul University Office for Teaching, Learning, and Assessment, accessed August 9, 2014, <http://condor.depaul.edu/tla/AssessmentMain.html>; "Academic Program Review," DePaul University Office of the President, accessed August 10, 2014, <http://pres.depaul.edu/aprc/>.

8. "Teaching and Learning Certificate Program," DePaul University Office for Teaching, Learning, and Assessment, accessed August 9, 2014, <http://tlcp.depaultla.org>; Richardson Library Information Commons, DePaul University Library, accessed August 9, 2014, <http://libguides.depaul.edu/infocommons>.

9. "Assessment," DePaul University Division of Student Affairs, accessed August 10, 2014, <http://studentaffairs.depaul.edu/assessment.html>.

10. Scott Walter and Michele Eodice, "Meeting the Student Learning Imperative: Supporting and Sustaining Collaboration Between Academic Libraries and Student Services Programs," *Re-*

*search Strategies* 20, no. 4 (2005): 219-225.

11. "Office of Multicultural Student Success," DePaul University Division of Student Affairs, accessed August 10, 2014, <http://studentaffairs.depaul.edu/omss/>; Richardson Library Learning Commons, DePaul University Library, accessed August 10, 2014, <http://libguides.depaul.edu/learningcommons>.

12. "DePaul's Executive Retention Group," DePaul University Division of Enrollment Management and Marketing, accessed August 9, 2014, <http://offices.depaul.edu/emm/student-retention/executive-retention-group/Pages/default.aspx>.

13. "Higher Learning Commission Initiative: Foundations for Success 2013-2016," DePaul University Division of Enrollment Management and Marketing, accessed August 10, 2014, <http://offices.depaul.edu/emm/student-retention/executive-retention-group/Pages/initiatives.aspx>; "BlueStar," DePaul University Teaching Commons, accessed August 11, 2014, <http://teachingcommons.depaul.edu/technology/bluestar.html>.

14. "Ithaka S+R Local Surveys," Ithaka S+R, accessed August 10, 2014, [www.sr.ithaka.org/content/ithaka-sr-local-surveys](http://www.sr.ithaka.org/content/ithaka-sr-local-surveys).

15. Meredith Gorran Farkas, Lisa Janicke Hinchliffe, and Amy Harris Houk, "Bridges and Barriers: Factors Influencing a Culture of Assessment in Academic Libraries," *College & Research Libraries* 76 (forthcoming), accessed August 10, 2014, <http://crl.acrl.org/content/early/2014/03/03/crl14-601.full.pdf+html>. *ZZ*

---

*("Internet resources," cont. from page 495)*

• **Linked Data: Evolving the Web into a Global Data Space.** This online book published in 2011 provides a broad look at Linked Data, exploring how the technologies of Linked Data (such as RDF/XML, OWL, SKOS, and SPARQL) come together to make linked open data work. Although technical in many places, the book can serve as a good textbook for information professionals who want to understand the underlying infrastructure that makes linked open data possible. Access: <http://linkeddatabook.com/>. *ZZ*