



Linked Data in a Nutshell: A Starter Kit of Selected Annotated Bibliography and Resources for Academic Librarians: Part One

Weiling Liu

Abstract:

It has been a decade since Tim Berners-Lee coined Linked Data in 2006. More and more Linked Data datasets have been made available for information retrieval on the Web. It is essential for librarians, especially academic librarians, to keep up with the state of Linked Data. There is so much information about Linked Data that one may wonder where to begin when they want to join the Linked Data community. With this in mind, the author compiled this annotated bibliography as a starter kit. Due to the many resources available, this list focuses on literature in English only and of specific projects, case studies, research studies, and tools that may be helpful to academic librarians, in addition to the overview of Linked Data concept and the current state of Linked Data evolution and adoption.

To cite this article:

Liu, W. (2018). Linked Data in a nutshell: a starter kit of selected annotated bibliography and resources for academic librarians: part one. *International Journal of Librarianship*, 3(1), 36-53. doi: <https://doi.org/10.23974/ijol.2018.vol3.1.47>

To submit your article to this journal:

Go to <http://ojs.calaijol.org/index.php/ijol/about/submissions>

Linked Data in a Nutshell: A Starter Kit of Selected Annotated Bibliography and Resources for Academic Librarians: Part One

Weiling Liu

University of Louisville, Louisville, Kentucky, USA

ABSTRACT

It has been a decade since Tim Berners-Lee coined Linked Data in 2006. More and more Linked Data datasets have been made available for information retrieval on the Web. It is essential for librarians, especially academic librarians, to keep up with the state of Linked Data. There is so much information about Linked Data that one may wonder where to begin when they want to join the Linked Data community. With this in mind, the author compiled this annotated bibliography as a starter kit. Due to the many resources available, this list focuses on literature in English only and of specific projects, case studies, research studies, and tools that may be helpful to academic librarians, in addition to the overview of Linked Data concept and the current state of Linked Data evolution and adoption.

Keywords: Linked Data, Linked Open Data, Library Linked Data, Semantic Web, Academic Library, Bibliography

INTRODUCTION

Linked Data is an approach to publishing data using URIs (Uniform Resource Identifiers) on the Web that can be shared and re-used by humans and computers. Linked Data are empowered by semantic technologies such as RDF (Resource Description Framework), SPARQL (a query language for RDF), OWL (Web Ontology Language), and SKOS (Simple Knowledge Organization System). Linked Data offers opportunities to extend the Web (the Web of Documents) (Linked Data Glossary, n.d.) towards a Semantic Web which is also called “the Web of Data” (Wood, Zaidman, Ruth, & Hausenblas, 2014, p. 3). The impact of Linked Data has shown on users’ search experiences and expectations on the Web (e.g., Google, Facebook).

Tim Berners-Lee first coined the term, Linked Data, in 2006. Since then, Linked Data datasets have grown considerably worldwide although there are still various challenges and issues

to be resolved, and even questions about its future. In the United States, some large-scale Linked Data datasets have been available or notable projects have been launched for years. For example, BIBFRAME (Bibliographic Framework Transition Initiative) by Library of Congress, VIAF (Virtual International Authority File) by Online Computer Library Center (OCLC), and Medical Subject Headings (MeSH) RDF by the National Library of Medicine. Two examples of new high-profile projects started after the W3C (World Wide Web Consortium) Library Linked Data Incubator Group's 2011 final report include BIBFLOW and Linked Data for Libraries (LD4L).

The library, archive, and museum (LAM) community are encouraged to bring their resources into the Semantic Web environment via Linked Data (Coyle, 2013, p. 57). They are strongly encouraged "to become more knowledgeable and attuned to the development and rollout of BIBFRAME and how it fits within libraries and the larger Linked Data sphere" (Library of Congress, 2016). Also, as Mitchell suggests, it is essential to get more library staff up to speed with the current state of Linked Data and the future changes influenced by Linked Data (Mitchell, 2016). Motivated by these statements, the author has created this annotated bibliographic list as a starter kit, especially for those working in an academic setting.

There is a lot of literature documenting Linked Data's development, such as case studies, experiments, and research studies around the world. This kit focuses on literature and sources in English only and of specific projects, case studies, research studies, and tools. The kit is divided into two parts: Part One focuses on Linked Data concepts, current state, and case studies of Linked Data evolution and adoption in a university setting. Part Two focuses on research studies of Linked Data vocabularies, Library Linked Data, novel approaches and tools used to address issues and challenges related to Linked Data implementation.

Included in Part One are references to five books, 17 journal articles, 14 websites, three reports, and several web-based articles. Through these resources, readers will find answers to some fundamental questions such as: a) What is Linked Data (LD)?; b) Linked Open Data (LOD), and Library Linked Data or Linked Library Data (LLD)?; c) How do they differ from each other?; What are the benefits of using Linked Data? and; d) How is it implemented and what is the current state of Linked Data adoption in the LAM community? All the annotated resources in the sections below, except those in the Website section, are listed in the References. For readers to easily follow the Linked Data development, the list in each section is ordered in ascending chronological order by year then by author. In the description and annotation of this kit, the term, Linked Data, will be used as a whole including LLD and LOD unless it is specified in a citation. The author hopes that through Part One, readers will come to a basic understanding of Linked Data, an overview of Linked Data publishing and consumption, and the current state of Linked Data development in the LAM community.

CONCEPTS, BENEFITS AND CURRENT STATE

In this section, readers will find selected resources about Linked Data, and technologies used to implement it and the current state of Linked Data evolution and adoption. For a quick start, readers may begin with the blog article by Jennifer Lohmann. Readers will see the original articles in which the concepts of the Semantic Web, Linked Data and Linked Open Data were first proposed. Readers will find detailed explanations of Linked Data's potential benefits to the libraries, related issues, and recommendations for all library stakeholders (leaders, standard bodies, data and system designers, librarians, and archivists). Last, in this section, readers will find some experts' views on the future of Linked Data.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*. Retrieved from http://www-sop.inria.fr/acacia/fabien/lecture/licence_travaux_etude2002/TheSemanticWeb/

The authors of this article describe an expected evolution of the current Web to a Semantic Web and define that Semantic Web is an extension of the existing Web. In Semantic Web, information is given well-defined meaning and able to be processed by computer software agents for the human users. In the article, the authors also describe how Semantic Web can be realized regarding expressing meaning, knowledge presentation, ontologies and software agents.

Berners-Lee, T. (2006, July 27). Linked data - Design issues. Retrieved from <https://www.w3.org/DesignIssues/LinkedData.html>

This article discusses the solutions to the problems of Semantic Web, which define Linked Data and the four principles of Linked Data. Later in 2010, Berners-Lee also added the information about Linked Open Data and the 5-star Linked Open Data standards. It is in this article that Berners-Lee, who proposed the Semantic Web in 2001, coined the term Linked Data.

Baker, T., Bermès, E., Coyle, K., Dunsire, G., Isaac, A., Murray, P., . . . Zeng, M. (2011). *Library Linked Data Incubator Group final report*. Retrieved from: <https://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/>

This final report is by W3C Library Linked Data Incubator Group which was chartered from May 2010 through August 2011 to help increase global interoperability of library data on the Web. Through the analysis of use case survey and inventory of available datasets, the report covers the benefits of Linked Data, current issues related to existing library data and Linked Data initiatives, and recommendations for the future. In the appendix, it also includes the survey and inventory results as well as descriptions of relevant technologies and semantic "Alignments" – links among vocabularies, metadata element sets, or datasets.

This report is an excellent starting point resource to get an overview of Library Linked Data. It includes definitions of Linked Data, Linked Open Data, and Library Linked Data. Although there have been many updates to the state of Linked Data development and adoption

over the past few years, the critical recommendations given in the report for all stakeholders remain significant.

Coyle, K. (2013). Linked Data: An evolution. *JLIS.It*, 4(1), 53-61. doi:10.4403/jlis.it-5443

Coyle briefly summarizes the characteristics of users and resources that libraries are facing nowadays. Coyle also highlights some library initiatives that follow Semantic Web standards and points out the importance of the library's role in today's information age: making library data available online where it will interact with existing and future information resources. Coyle believes that Linked Data can provide significant support to libraries to ensure their leadership role in the future.

Pesch, O., & Miller, E. (2016). Using BIBFRAME and Library Linked Data to solve real problems: An interview with Eric Miller of Zepheira. *The Serials Librarian*, 71(1), 1–8. doi:10.1080/0361526X.2016.1183159

This paper records the conversation between Pesch and Miller about the past, present, and future of BIBFRAME (Bibliographic Framework Initiative). Questions cover many aspects including libraries' missions and the role the Web plays in these mission statements, the basic premise behind the creation of BIBFRAME, to what extent Linked Data is useful now. Following the interview, Miller provides advice to libraries.

Mitchell, E. (2016, January 4). A Linked Data landscape: Critical decisions for data licensing, shared standards, and system design. *American Libraries*. Retrieved from <https://americanlibrariesmagazine.org/2016/01/04/linked-data-landscape/>

Mitchell gives a summary of the current state of Linked Data adoption in the LAM community. He suggests three focus points: education – training library staff for the upcoming changes; experimentation – exploring new systems with caution; and flexibility – monitoring standards as they are available.

Mitchell, E. T. (2016). Library linked data: Early activity and development [Monograph]. *Library technology reports*, 52(1), 5-33. Retrieved from <https://journals.ala.org/index.php/ltr/issue/view/534>

Mitchell gives an update on the broader state of Linked Data adoption focusing on the 2014 OCLC's surveys of Linked Data adoption and updates from national and international project teams. In the 2013 issue of Library Technology Reports (vol. 50, no. 5), Mitchell explored the fundamentals of Linked Data and its related problems as well as the state of adoption and technology use across the LAM community. Mitchell revisits this topic in this report, but is focusing on current trends, challenges, and opportunities.

For example, the emerged technical concepts in Linked Data such as TriG, N-Quads, and JSON-LD that may influence Linked Data adoption in the future. In Chapter 2, Mitchell illustrates the trends explored in Chapter 1 by examining a few representative projects and research initiatives in the Linked Data community. The projects include BIBFRAME, BIBFRAME Lite, Europeana, British Library and British Museum programs, and advances in OCLC's Linked Data projects. Mitchell also discusses some high-profile research projects and initiatives such as BIBFLOW and LD4L. The broad trends observed by Mitchell through the review of projects and literature include the following:

- an increasing interest in offering SPARQL endpoints as part of data publishing
- the distinction between discovery (end-user), access/service (developer/professional), and policy/rights (legal) perspectives in LD services
- the increasing need to bring together URI minting services and ensure that vocabulary adoption is done in a manageable way
- the discussion around comprehensive versus distributed standards
- the value of peer-to-peer metadata sharing and linking versus large or centralized sharing
- reconciliation and interoperability across metadata standards (Mitchell, 2016, p. 19)

In Chapter 3, Mitchell examines trends around specific approaches, application systems, vocabularies, standards, and tools used in the projects mentioned in Chapter 2. Mitchell concludes that the cases and studies explored in Chapter 2 indicate that “the LD LAM community is reaching a level of maturity that may be shaping next steps in Linked Data adoption toward production systems and permanent migration” (p. 22). Mitchell points out that Oslo Public Library's transformation to Linked Data-based services is an example of one project that has reached this goal.

In Chapter 4, Mitchell explores the emerging issues of Linked Data adoption in the LAM community such as data openness, standards compatibility and lack of support systems. Mitchell also reviews some critical broad questions and topics of future exploration, for example, the role of large-scale projects and the impact of Linked Data on cataloger work and notions of value in the future. Mitchell concludes that considerable progress has been seen in the past two years in building applications, vocabularies and best practices in the LAM community.

Zaino, J. (2016, November 29). 2017 Trends for Semantic Web and Semantic technologies. Retrieved from <http://www.dataversity.net/2017-predictions-semantic-web-semantic-technologies/>

Zaino summarizes the views of a few experts on 2017 trends of Semantic Web and Semantic Web technologies. Experts shared similar thoughts on the slow progress in the broader adoption of Semantic Web, such as Linked Data. They believed that the future would be focusing on Artificial Intelligence (AI). However, Semantic Web remains essential to the development of AI, as stated: “never doubt one thing: That is, that the Semantic Web, and Semantic technologies

and techniques, have and will continue to propel the industry forward.” The experts include Amit Sheth, Michael Bergman, and David Wood. Sheth is the executive director of Kno.e.sis—the Ohio Center of Excellence in Knowledge-enabled Computing, who is also an educator, researcher, and entrepreneur. Bergman, CEO of Structured Dynamics, is a co-founder of knowledge-based Artificial Intelligence startup Cognonto. Wood, an author of books including *Linking Enterprise Data*, is CTO of 3 Round Stones and Director of Technology at Ephox + TinyMCE.

Lohmann, J. (2017, April 19). *Linked Data definitions: Bringing clarity to Linked Data terms and buzz words*. [Blog post]. Retrieved from <https://medium.com/library-link-network/linked-data-definitions-75d6e199d3ad>

The article starts with the concept of the MARC record and illustrates Linked Data usage with two screenshots of web search results supported by Linked Data. At the end of the article, Lohmann gives a brief description of Library.Link, a network of Linked Data contributed by libraries. This article explains the Linked Data concept in a non-technical way and seems to be a quick way to understand the concept of Linked Data.

CASE STUDIES, EXPERIMENTS, AND PROJECTS

From the resources given above, the readers may have learned the concepts of Linked Data, Linked Open Data and Library Linked Data as well as Linked Data benefits to the LAM community. In this section, readers will find literature about Linked Data case studies, experiments, and projects. Except for one published in 2009, the resources selected in this section are within the past five years. The types of datasets used in the projects or initiatives, experiments, and case studies cover many areas in an academic setting, such as archives, legal documents, theses and dissertations, medical databases, and music databases. Involved systems include integrated library system, digital collection, discovery, and clinical decision support system.

Malmsten, M. (2009). *Exposing library data as Linked Data*. Retrieved from https://www.researchgate.net/publication/228391527_Exposing_Library_Data_as_Linked_Data

This article gives an overview of the Linked Data implementation with LIBRIS (<http://libris.kb.se>), the Swedish Union Catalogue. It describes the rationales behind this application as well as steps needed to convert the catalog to Linked Data. It also explains a minimal API (Application Program Interface) for exporting bibliographic data and relations from an Integrated Library System.

Lampert, C., & Southwick, S. (2013). *Leading to linking: Introducing Linked Data to academic library digital collections*. *Journal of Library Metadata*, 13(2/3), 230–253. doi:10.1080/19386389.2013.82609

This article presents an overview of a project that transforms metadata of an academic library's digital collections into Linked Data. It begins with an introduction of Linked Data concepts within the context of digital collections. The article covers the rationale for the project, technologies used for the data transformation, publishing and management, lessons learned, and the approach the project team took to learn about Linked Data concepts.

Thompson, K., & Richard, J. (2013). Moving our data to the Semantic Web: Leveraging a content management system to create the linked open library. *Journal of Library Metadata*, 13(2–3), 290–309. doi:10.1080/19386389.2013.828551

This article describes two pilot projects of Smithsonian Libraries. One uses Drupal (version 7), a content management system, to publish bibliographic data taken from the library catalog and present it as Linked Open Data. The other creates Linked Open Data from a much-cited botanical reference work, Taxonomic Literature II (TL-2).

Smithsonian Libraries successfully published their 4,000 digital objects as Linked Open Data and converted TL-2 to Linked Open Data with "4-star" quality scale. In the lessons learned section, the authors point out that the most challenging aspect of publishing Linked Open Data with "5-star" [see Berners-Lee's (2006) article in the previous section for Linked Open Data definition] is to identify correct links in other data sources on the Web. It is also difficult to create Linked Open Data links especially when there are few or no common identifiers in the local dataset.

Hanson, E. M. (2014). A beginner's guide to creating Library Linked Data: Lessons from NCSU's Organization Name Linked Data project. *Serials Review*, 40(4), 251–258. doi:10.1080/00987913.2014.975887

In this article, Hanson describes the Organization Name Linked Data (ONLD) project at the North Carolina State University Libraries (NCSU), a leading partner of GOKb project. This article also covers project planning in addition to exploring best practices for data clean up, conversion, and publishing linked data. In conclusion, Hanson states that the NCSU ONLD project has proved to be an excellent Linked Data pilot.

Hanson points out that the project was completed in a relatively short time frame with a small team although the creation of links between locally created datasets and externally linked datasets was manually done. Tools and workflows developed during the project can be used for future projects. Hanson also encourages librarians to get involved, especially those who work with serials.

Mak, L., Higgins, D., Collie, A., & Nicholson, S. (2014). Enabling and integrating ETD repositories through Linked Data. *Library Management*, 35(4/5), 284–292. doi:10.1108/LM-08-2013-0075

In the article, the authors describe the process of developing a Linked Data model that uses Electronic Theses and Dissertation (ETD) data. Through the analysis of graph data, the article illustrates a new perspective in examining the academic research landscape which describes the relationships between departments, faculty members, themes, and dissertation authors.

Antelman, K., & Wilson, K. (2015). The Global Open Knowledgebase (GOKb): Open Linked Data supporting electronic resources management and scholarly communication. *Insights*, 28(1), 42-50. doi:10.1629/uksg.217

This article gives a detailed overview of GOKb (<http://gokb.org>) and its potential benefits. GOKb is a repository of open data related to library e-resources. GOKb was initially a collaboration between Quali OLE (the Open Library Environment) partners and JISC of the United Kingdom. GOKb now has attracted others from all over the world. The article is based on a breakout session held at the 38th UKSG Annual Conference in Glasgow on March-April 2015.

Bushman, B., Anderson, D., & Fu, G. (2015). Transforming the Medical Subject Headings into Linked Data: Creating the authorized version of MeSH in RDF. *Journal of Library Metadata*, 15(3-4), 157-176. doi:10.1080/19386389.2015.1099967

This article reviews the project initiated by National Library of Medicine (NLM) that converts the Medical Subject Headings from XML to RDF. It also examines the collaborative process, the technical and organizational issues, and the future of Linked Data at the library. The authors point out that this MeSH RDF pilot has brought together representatives from many areas of NLM. It has dramatically improved the knowledge sharing across the library about Linked Data along with related aspects such as data modeling, application development, version control, and workflow management. It is only a beginning, and there are still many unsolved challenges, for example, the publication of MeSH RDF is not yet real Linked Data, and the decision on data modeling still needs more discussions.

Lafia, S., Jablonski, J., Kuhn, W., Cooley, S., & Medrano, F. A. (2016). Spatial discovery and the research library. *Transactions in GIS*, 20(3), 399-412. doi:10.1111/tgis.12235

This article presents a proof-of-concept model developed in a joint research project by the Library and the Center for Spatial Studies, both at the University of California, Santa Barbara (UCSB). This model uses Linked Data solution in which spatial and non-spatial research data can be shared in a self-deposit environment like ArcGIS Online. Through this model, the authors demonstrate how academic libraries can increase the discoverability and reusability of research publications and data across disciplinary boundaries by linking publications stored in a library repository to data hosted on an external system like Esri Open Data.

van Ballegooie, M., Borie, J., & Senior, A. (2017). The Canadian Linked Data Initiative: Charting a path to a Linked Data future. *The Serials Librarian*, 72(1-4), 207-213. doi:10.1080/0361526X.2017.1292751

This article is a preliminary report on the progress of CLDI (Canadian Linked Data Initiative), which is a collaboration between Library and Archives Canada, Bibliothèque et Archives nationales du Québec, Canadiana.org, and five of Canada's largest research libraries. This report covers the work of CLDI in five critical areas of activity: digital collections, education and training, legacy metadata enhancement, grant funding, and the adaptation and evaluation of BIBFRAME tools. This article reports on the first nine months of the partnership since September 2015 when the CLDI was formed and also outlines the new opportunities and challenges in this collaborative project. For example, the development of cross-country, cross-institutional collaboration is one of the most immediate opportunities. A problem is how to bring individual library units onboard when both human and funding resources are scarce.

Marco-Ruiz, L., Pedrinaci, C., Maldonado, J. A., Panziera, L., Chen, R., & Bellika, J. G. (2016). Publication, discovery and interoperability of Clinical Decision Support systems: A Linked Data approach. *Journal of Biomedical Informatics*, 62, 243–264. doi:10.1016/j.jbi.2016.07.011

This article is a study in which Linked Data principles are used to develop methods and models to enhance Clinical Decision Support (CDS) services. The three areas the study focused on include developing ways to define the CDS service semantics, discovering services based on such methods, and analyzing their metadata, functionality and data models. The study results show that Linked Data makes it possible to integrate CDS services under shared Linked Knowledge Bases and to use intelligent queries for their discovery in a health network. Linked Data also enables the possibility to analyze services based on rich unambiguous semantics that powers the interoperability and reuse of CDS services.

Page, K. R., Bechhofer, S., Fazekas, G., Weigl, D. M., & Wilmering, T. (2017). Realising a layered digital library: Exploration and analysis of the Live Music Archive through Linked Data. In *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)* (pp. 1–10). doi:10.1109/JCDL.2017.7991563

This article illustrates how CALMA (Computational Analysis of the Live Music Archive), a digital library layered over the Live Music Archive, was built using a Linked Data approach. The Live Music Archive is part of an online resource, Internet Archive, which provides access to a sizeable community-contributed collection of live recordings. Through concrete examples and exploratory investigation, the authors demonstrate the benefits of CALMA in the capability and flexibility it can bring to scholarly users.

As summarized in the article, there have been some previous successful implementations of applying Semantic Web technologies in the context of digital music collections with similar research focuses. However, there are some limitations of those applications. CALMA is described as an implemented example of Linked Data of practical use which exhibits as a better solution regarding information consumption, enrichment, and reuse.

Szöke, Á., Föhrécz, A., Kőrösi, G., & Strausz, G. (2013). Versioned linking of semantic enrichment of legal documents. *Artificial Intelligence and Law*, 21(4), 485–519. doi:10.1007/s10506-013-9145-z

This article illustrates the benefits of Linked Data in improving access to legal information on the Web. The proposed semantic enrichment approach presented in this article is a combination of a knowledge base-based enrichment method of legal texts and a linking method. It is intended to address three issues: 1) the ambiguousness of legal texts, 2) utilizing the application of legal reasoners, and 3) change management of legal texts. The review of current approaches in the field, such as Akoma Ntoso and MetaLex, show that solutions to these issues were not found or only partially handled. However, the new semantic enrichment approach, as the case study shows, can address these problems. The article provides useful information on background and the new enrichment methodology, great detailed explanations on the identification and the versioning of legal resources, and a good overview of the linking approach and the proof-of-concept experiment.

RESOURCE FOR IMPLEMENTATION

This section includes selected bibliographies of literature for publishing and consuming Linked Data. Readers will learn more details about Linked Data, Linked Open Data, Library Linked Data and tools for use in the development as well as various Linked Data related resources. Readers may find information about consumption of Linked Data in many of the resources on this list, but there is an article reviewing Linked Data browsers and exploratory search system.

Heath, T., & Bizer, C. (2011). *Linked Data: Evolving the web into a global data space. Synthesis lectures on the semantic web: Theory and technology*, 1:1. San Rafael, CA: Morgan & Claypool. Retrieved from <http://linkeddatatoolkit.com/editions/1.0>

This book provides an overview of the principles of Linked Data and relevant aspects of the Web of Data. It also provides guidance and best practices on two main topics – the publication and consumption of Linked Data. There are seven chapters in the book. The first three chapters cover the overview of Linked Data principles and the Web of Data. Chapter 4 and 5 are about architectural approaches to publishing Linked Data, methods for automated linking of data sets, and how to choose URIs and vocabularies to identify and describe resources as well as debugging approaches for Linked Data deployment. Chapter 6 gives an overview of existing Linked Data applications and tools, and Chapter 7 is the summary and outlook. Through this book, “[r]eaders can expect to gain a rich technical understanding of Linked Data fundamentals, as the basis for application development, research or further study” (Heath & Bizer, 2011).

Hooland, S., & Verborgh, R. (2014). *Linked Data for libraries, archives and museums: How to clean, link and publish your metadata*. Chicago, IL: Neal-Schuman

In the book, the authors explore the evolution of metadata and implementation of Linked Data in theory and practice, with a focus on the common metadata shared by the LAM community through the discussions and illustrations of case studies.

This handbook is very informative and well-structured with practical hands-on examples. It is an excellent resource for those who are interested in learning more about Linked Data and wishing to get started with their institution's Linked Data projects. There are seven chapters, and five of them are on specific topics: modeling, cleaning, reconciling, enriching, and publishing. Each of these sections gives its goal, target audience, conceptual insights and practical skills as well as case studies. The book also provides links to websites for downloading the metadata used in each case study.

Wood, D., Zaidman, M., Ruth, L., & Hausenblas, M. (2014). *Linked Data: Structured data on the Web*. Shelter Island, NY: Manning.

This book presents the Linked Data concepts in context, covering techniques for how to publish Linked Data and optimize Web pages with Linked Data for searching as well as how to consume Linked Data. It brings the contents to readers in step-by-step work through examples of increasing complexity. The book has four parts and eleven chapters. The book also provides links to source codes used in the book. Appendices include supplemental information such as a reference to development environment setups of the tools used in the book. The intended audience is Web developers who are interested in publishing and consuming Linked Data. However, it is written in such an easy to understand manner; it does not require any previous knowledge about Linked Data or Semantic Web technologies but a basic familiarity with Web technologies such as HTML, HTTP, and URIs.

Godby, C., Wang, S., & Mixter, J. (2015). *Library linked data in the cloud: OCLC's experiments with new models of resource description*. San Rafael, CA: Morgan & Claypool.

This book describes OCLC's experiments with Library Linked Data and its contributions to the transformation of legacy data from a Web of Documents to a Web of Data. It focuses on modeling bibliographic descriptions as Linked Data not only in theory but also with technical details and practical applications. The main topics covered in chapters are the Semantic Web and library standards, modeling library authority files and creative works, discovering creative works, using text mining for entity identification, and the Library Linked Data cloud. At the end of the book, the authors address a few lessons learned and challenges confronted. The book is intended for those who are interested in modeling bibliographic descriptions as Linked Data, but beginners may get an overview of what OCLC has been working on regarding Linked Data.

Jacksi, K., Dimililer, N., & Zeebaree, S. R. M. (2016). State of the art exploration systems for Linked Data: A review. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 7(11). doi:10.14569/IJACSA.2016.071120

The article gives a review of Semantic Web technology and an overview of the search strategies. The article provides the survey results of state-of-the-art Linked Data Browsers (LDBs) and Exploratory Search Systems (ESSs) based on Linked Open Data. The article also compares each of the LDBs and ESSs concerning features such as algorithms, data presentations, and explanations.

As the use of Linked Open Data has notably increased on the Web, the challenge for users becomes more apparent, and a lot of LDBs have been developed to address issues. The detailed review of these LDBs and ESSs is beneficial for librarians as a Linked Data consumer as well as a contributor.

Jones, E., & Seikel, M. (Eds.). (2016). *Linked data for cultural heritage*. Chicago, IL: American Library Association.

In the book, the editors present six contributions by an authoritative group of authors including librarians, technology experts, and project managers who implemented Linked Data projects across the LAM community, to help readers understand Linked Data concepts and its purpose. The book begins with a brief overview of the basic description of linked data and its promises and challenges. Individual chapters cover different topics including a survey of several Linked Open Data projects around the world, controlled vocabularies, and their purpose, best practices regarding creating library thesauri, metadata element sets, OCLC's experiment with Schema.org, and BIBFRAME data model. It is "[r]ecommended for readers interested in LIS and the history and methods of disseminating information in virtual environments" (Harris, 2016, p.108). More specifically, readers "will find current perspectives on such questions as what linked data can or cannot (yet) do, what kinds of tools exist to assist the conversion, what level of human intervention will be needed, why are controlled vocabularies needed, and how can they be found and selected" (Ojennus, 2017, p.174).

Coyle, K. (2012). *Linked Data Tools: Connecting on the Web* [Monograph]. *Library Technology Reports*, 48(4), 6-46. Retrieved from <https://journals.ala.org/index.php/ltr/issue/view/183>

Coyle presents an overview of current library data activities and introduces Semantic Web concepts in the context of library data, including primary components of Linked Data. Coyle discusses metadata elements, the building blocks of metadata scheme as well as controlled vocabularies for Linked Data. Coyle also covers in the last two chapters various tools and resources on Linked Data for both beginners and advanced users, Linked Data implementation, and related communities.

WEBSITES

This section aims to provide readers with a quick reference list of selected website resources. Linked Data standards, library linked datasets or Linked Data services, notable library Linked Data

projects or initiatives in the United States, tools, and subject conferences. Most of these resources are mentioned in the literature in section 1-3. However, the author identifies them here for readers' convenience because they are not only good for keeping up-to-date on Linked Data development in the libraries but also essential for implementation.

The American Art Collaborative: Linked Open Data Initiative
(<http://americanartcollaborative.org/>)

The American Art Collaborative (AAC) is a consortium of 14 art museums in the United States. The Collaborative believes that Linked Open Data offers the productive potential to increase the understanding of art and is committed to building a network of practitioners, sharing information with the library, archives, and museum community, and establishing a critical mass of Linked Open Data on the Web. The website includes various links to related information and resources such as "Linked Data and Tools," "AAC Browse App Demo," and "The Linked Art Data Model."

BIBFLOW (<https://bibflow.library.ucdavis.edu/>)

BIBFLOW, funded by the Institute of Museum and Library Services, is a two-year project of the University of California, Davis' University Library and Zepheira. As its official title "Reinventing Cataloging: Models for the Future of Library Operations" indicates, this project is investigating the future of library cataloging and related workflows, in light of modern technology infrastructures such as the Web, new data models, and formats such as RDA (Resource Description and Access) and BIBFRAME. The roadmap of this project is available on the website.

DBpedia (<http://dbpedia.org/>)

DBpedia is an open community project. The structured information found in DBpedia is extracted from Wikipedia with crowd-sourced community effort and available on the Web. DBpedia makes it possible for search engines like Google to use sophisticated queries against Wikipedia, and to link the different datasets on the Web to Wikipedia data.

JSON for Linked Data (<https://json-ld.org/>)

JavaScript Object Notation (JSON) for Linked Data (JSON-LD), based on the already successful JSON format, is a lightweight Linked Data format which is easy for both machines to parse and generate and humans to read and write. This site includes some popular programming environments and resources for developers. It also provides a playground not only helpful to developers but also general users who are interested in learning JSON-LN.

Library.Link Network (<http://library.link>)

This website is the home of the Library.Link Network (LLN). LLN was created as a collaboration by Zepheira and library industry partners: Atlas Systems, EBSCO Novelist, Innovative Interfaces, and SirsiDynix. Library.Link provides a platform on behalf of libraries for libraries to publish their content on the Web using Linked Data standards and Semantic technologies. “Link Domain Data, Filters and Services for University of Manitoba Libraries (<http://link.lib.umanitoba.ca/data.html>)” is an example of library data published using Library.Link.

The Library of Congress: BIBFRAME (Bibliographic Framework Initiative)
(<https://www.loc.gov/bibframe/>)

BIBFRAME is an initiative by the Library of Congress. A primary focus of this effort is to explore a way for the transition from the MARC 21 formats to BIBFRAME, a data model that makes bibliographic data more useful both within and outside the library community using Linked Data principles.

The Library of Congress Linked Data Service: Authorities and Vocabularies.
(<http://id.loc.gov/>)

Influenced by Linked Data, the Library of Congress’ Linked Data Service provides commonly found standards and vocabularies published by the Library of Congress that are accessible by both humans and machines. The website includes data values and the controlled vocabularies such as Library of Congress Subject Headings.

LD4L: Linked Data for Libraries (<https://ld4l.org/>)

The Linked Data for Libraries (LD4L) website, as described on the site, is a gateway to Linked Data projects for library communities. LD4L website contains information on Linked Data for Production (LD4P) (2016-2018) and LD4L Labs (2016-2018) as well as LD4L (2014–2016). LD4P focuses on production workflows based on Linked Open Data. LD4L Labs focuses on developing tools and providing support to pilot solutions that can be in production at research libraries in the future. LD4L projects are all collaborations and funded by the Andrew W. Mellon Foundation.

National Library of Medicine: Medical Subject Headings RDF (<http://id.nlm.nih.gov/mesh>)

This website is the home for MeSH RDF. MeSH is a Linked Data representation of the MeSH biomedical vocabulary produced by the National Library of Medicine. Resources offered on MeSH RDF includes a downloadable SPARQL query editor, a RESTful (REST - Representational State Transfer) interface for retrieving MeSH data as well as a file in RDF N-Triples format, a SPARQL endpoint API, and some sample queries.

OCLC Linked Data (<http://www.oclc.org/developer/develop/linked-data.en.html>)

This page summarizes OCLC's Linked Data development including datasets, ongoing projects and products in development. For example, the Dewey Decimal Classification (DDC), the Virtual International Authorities File (VIAF) and Faceted Application of Subject Terminology (FAST) that have been published as Linked Data. WorldCat Linked Data is an ongoing experimental project.

OpenRefine (<http://openrefine.org/>)

As described on the site, OpenRefine (a free open source and formerly Google Refine) is a powerful tool for working with messy data. It can be used for cleaning and transforming data from one format to another. It is now a community site and has been maintained by volunteers since 2012 when Google stopped supporting this project.

Schema.org (<http://schema.org/>)

Schema.org, an initiative launched by Bing, Google, and Yahoo in 2011 and later joined by Yandex, is a collaborative, community activity. It intends to use the schema.org vocabulary along with the Microdata, RDFa (Resource Description Framework in Attributes), or JSON-LD formats to markup website content with metadata about itself. OCLC has been working with Schema.org for years on its Linked Data projects such as WorldCat.org by appending Schema.org descriptive mark-up to WorldCat.org pages.

SWIB - Semantic Web in Libraries (<http://swib.org>)

This website is for Semantic Web in Libraries conferences. According to the information available on the site, SWIB conference is an annual conference, started in 2009, and is organized by ZBW - German National Library of Economics / Leibniz Information Centre for Economics and the North Rhine-Westphalian Library Service Centre (hbz). Since 2012, the conferences are in English. As stated on the site, the conference focuses on Linked Open Data in libraries and related organizations and is oriented not only for IT professionals or developers, but also librarians and researchers from all over the world: "WIB revolve around opening data, linking data and creating tools and software for Linked Open Data production scenarios. These areas of focus are supplemented by presentations of research projects in applied sciences, industry applications and Linked Open Data activities in other areas" (SWIB17, n.d.). Previous conference presentations are accessible from the History tab of the site.

W3C (World Wide Web Consortium) (<https://www.w3.org>)

The site is the official website of World Wide Web Consortium. It contains all the information about Linked Data and related standards and specifications (<https://www.w3.org/standards/semanticweb/data>) and resources for tools such as SPARQL (<https://www.w3.org/TR/rdf-sparql-query/>) and wiki (<https://www.w3.org/2001/sw/wiki/Tools>).

The wiki on tools is on W3C, but the community maintains its content. It includes a collection of references to complete development environments, editors, libraries or modules for various programming languages, specialized browsers, etc.

SUMMARY

Linked Data, explained as the Web of Data, will significantly benefit information search and discovery. The transition of library data from the format for the Web of Documents to the Web of Data will impact not only users' search experiences but also library's future workflows. This starter kit is expected to give academic librarians a good starting point to learn more about Linked Data. This unique list includes the references to original articles and resources about Linked Data standards and framework; examples of practices covering most domains of an academic library such as art, medical, and music; and different types of data: bibliographic data, and local metadata for archives, digital collections, and ETD. The list also has references to books on how to implement Linked Data as well as websites of standards, notable projects or initiatives, and tools. As a starting point, the focus of Part One is primarily on fundamental theory and practices for publishing and consuming Linked Data. Other aspects such as data management, specific applications, and assessment will be included in Part Two.

References

- Antelman, K., & Wilson, K. (2015). The Global Open Knowledgebase (GOKb): Open linked data supporting electronic resources management and scholarly communication. *Insights*, 28(1), 42-50. doi:10.1629/uksg.217
- Baker, T., Bermès, E., Coyle, K., Dunsire, G., Isaac, A., Murray, P., . . . Zeng, M. (2011). *Library Linked Data Incubator Group final report*. Retrieved from W3C Incubator. World Wide Web Consortium website: <https://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/>
- Berners-Lee, T. (2006, July 27). *Linked Data - Design issues*. Retrieved from <https://www.w3.org/DesignIssues/LinkedData.html>
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*. Retrieved from http://www-sop.inria.fr/acacia/fabien/lecture/licence_travaux_etude2002/TheSemanticWeb/
- Bushman, B., Anderson, D., & Fu, G. (2015). Transforming the Medical Subject Headings into Linked Data: Creating the authorized version of MeSH in RDF. *Journal of Library Metadata*, 15(3-4), 157-176. doi:10.1080/19386389.2015.1099967
- Cole, T. W., Han, M., Weathers, W. F., & Joyner, E. (2013). Library marc records into linked open data: Challenges and opportunities. *Journal of Library Metadata*, 13(2/3), 163-196. doi:10.1080/19386389.2013.826074
- Coyle, K. (2012). Linked Data tools: Connecting on the Web [Monograph]. *Library Technology Reports*, 48(4), 6-46. Retrieved from: <https://journals.ala.org/index.php/ltr/issue/view/183>

- Coyle, K. (2013). Linked Data: An evolution. *JLIS.It*, 4(1), 53. doi:10.4403/jlis.it-5443
- Godby, C., Wang, S., & Mixter, J. (2015). *Library linked data in the cloud: OCLC's experiments with new models of resource description*. San Rafael, CA: Morgan & Claypool. doi:10.2200/S00620ED1V01Y201412WBE012
- Gracy, K. F., Zeng, M. L., & Skirvin, L. (2013). Exploring methods to improve access to music resources by aligning library data with linked data: A report of methodologies and preliminary findings. *Journal of the American Society for Information Science and Technology*, 64(10), 2078.
- Hanson, E. M. (2014). A beginner's guide to creating Library Linked Data: Lessons from NCSU's Organization Name Linked Data project. *Serials Review*, 40(4), 251–258. doi:10.1080/00987913.2014.975887
- Harris, J. (2016). Linked data for cultural heritage. [Review of the book *Linked data for cultural heritage*, ed. by E. Jones & M. Seikel]. *Library Journal*, 141(20), 108.
- Heath, T., & Bizer, C. (2011). *Linked Data: Evolving the web into a global data space. Synthesis lectures on the semantic web: Theory and technology*, 1:1. San Rafael, CA: Morgan & Claypool. Retrieved from <http://linkeddatabook.com/editions/1.0>
- Hooland, S., Verborgh, R., & Neal-Schuman Publishers. (2014). *Linked Data for libraries, archives and museums: How to clean, link and publish your metadata* (U.S. edition. ed.). Chicago, IL: Neal-Schuman.
- Jacksi, K., Dimililer, N., & Zeebaree, S. R. M. (2016). State of the art exploration systems for Linked Data: A review. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 7(11). doi:10.14569/IJACSA.2016.071120
- Jones, E., & Seikel, M. (Eds.). (2016). *Linked data for cultural heritage*. Chicago, IL: American Library Association.
- Lafia, S., Jablonski, J., Kuhn, W., Cooley, S., & Medrano, F. A. (2016). Spatial discovery and the research library. *Transactions in GIS*, 20(3), 399–412. doi:10.1111/tgis.12235
- Lampert, C., & Southwick, S. (2013). Leading to linking: Introducing linked data to academic library digital collections. *Journal of Library Metadata*, 13(2/3), 230–253. doi:10.1080/19386389.2013.82609
- Library of Congress. (2016, November 22). BIBFRAME and PCC. Retrieved from <https://www.loc.gov/aba/pcc/bibframe/bibframe-and-pcc.html>
- Linked Data Glossary. (n.d.). Retrieved March 16, 2018, from <https://www.w3.org/TR/2013/NOTE-ld-glossary-20130627/#web-of-documents>
- Lohmann, J. (2017, April 19). *Linked Data definitions: Bringing clarity to Linked Data terms and buzz words*. [Blog post]. Retrieved from <https://medium.com/library-link-network/linked-data-definitions-75d6e199d3ad>
- Mak, L., Higgins, D., Collie, A., & Nicholson, S. (2014). Enabling and integrating ETD repositories through Linked Data. *Library Management*, 35(4/5), 284–292. doi:10.1108/LM-08-2013-0075
- Malmsten, M. (2009). *Exposing library data as Linked Data*. Retrieved from https://www.researchgate.net/publication/228391527_Exposing_Library_Data_as_Linked_Data

- Marco-Ruiz, L., Pedrinaci, C., Maldonado, J. A., Panziera, L., Chen, R., & Bellika, J. G. (2016). Publication, discovery and interoperability of Clinical Decision Support systems: A Linked Data approach. *Journal of Biomedical Informatics*, 62, 243–264. doi:10.1016/j.jbi.2016.07.011
- Mitchell, E. (2016, January 4). A Linked Data landscape: Critical decisions for data licensing, shared standards, and system design. *American Libraries*. Retrieved from <https://americanlibrariesmagazine.org/2016/01/04/linked-data-landscape/>
- Mitchell, E. T. (2016). Library Linked Data: Early Activity and Development [Monograph]. *Library Technology Reports*, 52(1), 5-33. Retrieved from <https://journals.ala.org/index.php/ltr/issue/view/534>
- Ojennus, P. (2017). [Review of the book *Linked data for cultural heritage*, ed. by E. Jones & M. Seikel]. *Library Resources & Technical Services*, 61(3), 173-174.
- Page, K. R., Bechhofer, S., Fazekas, G., Weigl, D. M., & Wilmering, T. (2017). Realising a layered digital library: Exploration and analysis of the Live Music Archive through Linked Data. In *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)* (pp. 1–10). doi:10.1109/JCDL.2017.7991563
- Pesch, O., & Miller, E. (2016). Using BIBFRAME and Library Linked Data to solve real problems: An interview with Eric Miller of Zepheira. *The Serials Librarian*, 71(1), 1–8. doi:10.1080/0361526X.2016.1183159
- SWIB17 - Semantic Web in Libraries. (n.d.). Retrieved from <http://swib.org/swib17/>
- Thompson, K., & Richard, J. (2013). Moving our data to the Semantic Web: Leveraging a content management system to create the linked open library. *Journal of Library Metadata*, 13(2–3), 290–309. doi:10.1080/19386389.2013.828551
- van Ballegooie, M. Borie, J., & Senior, A. (2017). The Canadian Linked Data Initiative: Charting a path to a Linked Data future. *The Serials Librarian*, 72(1–4), 207–213. doi:10.1080/0361526X.2017.1292751
- Wong, E. Y. (2017). LD4L: Linked Data for libraries. *Technical Services Quarterly*, 34(3), 332–333. doi:10.1080/07317131.2017.1321405
- Wood, D., Zaidman, M., Ruth, L., & Hausenblas, M. (2014). *Linked Data: Structured data on the Web*. Shelter Island, NY: Manning.
- Zaino, J. (2016, November 29). 2017 Trends for Semantic Web and semantic technologies. Retrieved from <http://www.dataversity.net/2017-predictions-semantic-web-semantic-technologies/>

About the author

Weiling Liu is Professor and Head of Office of Libraries Technology at University of Louisville, Louisville, Kentucky, USA.