

# **UCLA**

## **Librarian Publications**

### **Title**

UCLA Semantic Web and Linked Data LibGuide

### **Permalink**

<https://escholarship.org/uc/item/4jc1c4gn>

### **Author**

Super, Rhonda A., Ms.

### **Publication Date**

2022

### **Copyright Information**

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

# Semantic Web and Linked Data: Home

---

 libapps.s3.amazonaws.com/sites/705/guides/357266/backups/guide\_id\_357266\_1665509813.html

## Home

---

### Welcome!

---

The Semantic Web encompasses the technology that connects data from different sources across the Web as envisioned by Tim Berners-Lee and led by the World Wide Web Consortium (W3C). This Web of Data enables the linking of data sets across data silos on the Web by providing for machine-to-machine communication through the use of Linked Data. This Guide provides descriptions and links to resources used to implement this technology.

## Home

---

### NOTE:

The UCLA Semantic Web LibGuide was compiled and written by Rhonda Super. It began as a data page on Ms. Super's personal resource home page. Over a twenty year period, the Semantic Web resources listed on Rhonda's Resource Page developed into a stand alone LibGuide that served as a comprehensive resource for the Semantic Web and Linked Data community providing links to tools, best standards, instructional materials, use cases, vocabularies, and more.

The Guide was updated continuously through August 2022 using the SpringShare LibGuide platform as customized by the UCLA Library. Many of its resources provide a historical look at the development of Linked Data.

Ms. Super holds a BA in English and Government and an MA in Communications from Ohio University. She earned her MLIS from San Jose State University with a concentration in archives, rare books, and academic libraries. She earned a Certificate in XML and RDF Systems from the Library Juice Academy. Ms. Super was awarded scholarships to attend the California Rare Book School where she studied Rare Books for Scholars and Archivists, Descriptive Bibliography, and History of the Book: Nineteenth and Twentieth Centuries. Ms. Super was employed by the UCLA Library from 2007 to her retirement in 2022.

The final iteration of the Guide is deposited in the University of California eScholarship Open Access repository so the Linked Data community can continue to use it as a resource.

If you cite resources from this Guide, please check the original resource for copyright and citation requirements.

## Contents

---

Scroll down the page to access the topics listed below.

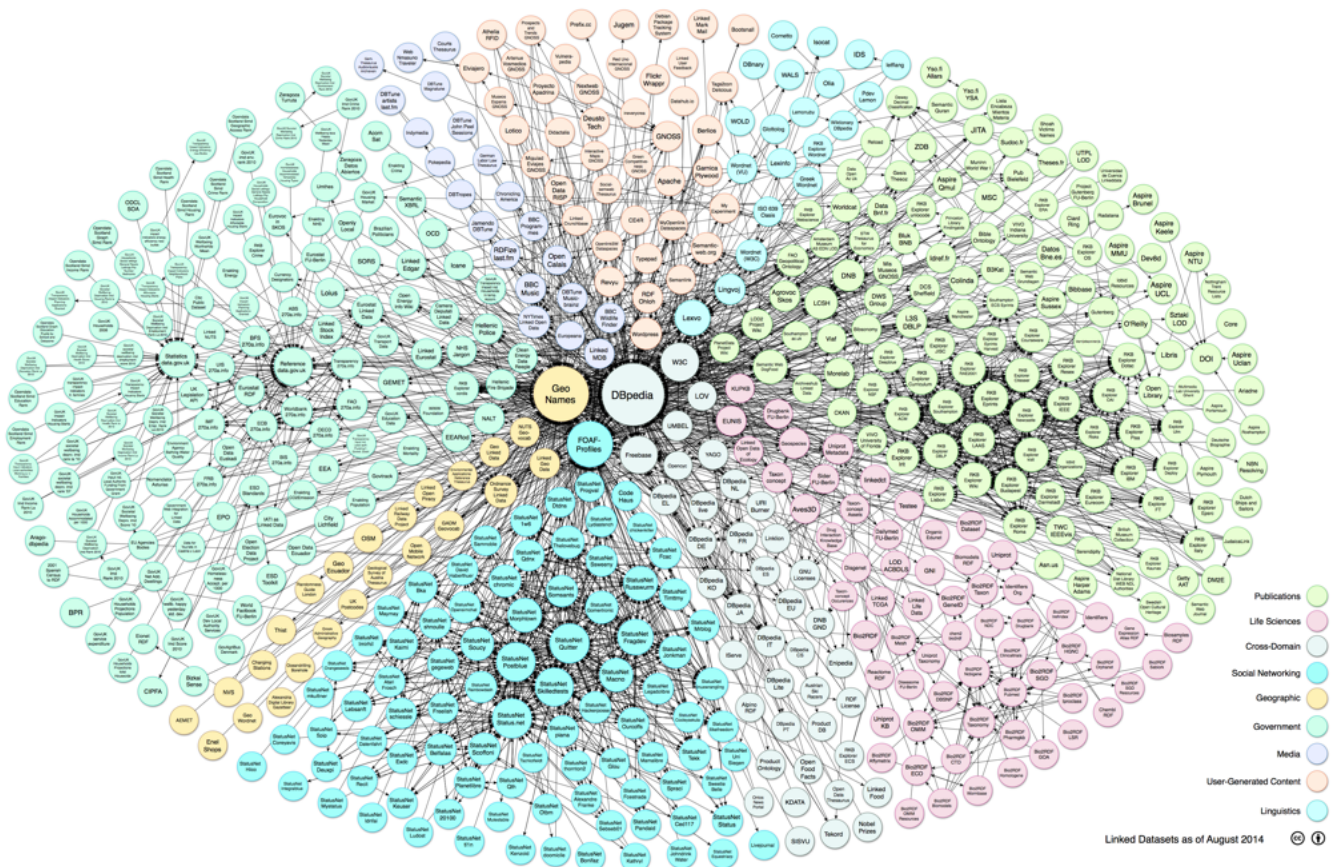
- Home
- Best Practices, Standards, and Metadata Application Profiles (MAPS)
- Blogs, Listservs, and Wikis
- Books
- Datasets
- Instructional Resources
- Journals, Articles, and Papers

- Semantic Web Services
- Semantic Web Tools
- SPARQL
- Vocabularies, Ontologies, and Frameworks
  - Ontologies and Frameworks
  - Registries, Portals, and Authorities
  - Schemas
  - Vocabularies
- Wikibase and Wikidata
  - WikiProjects
  - Wikidata Properties
  - Wikidata Tools
- Workshops and Projects
- Use Cases

## About the Semantic Web

The Semantic Web provides for the ability to semantically link relationships between Web resources, real world resources, and concepts through the use of Linked Data enabled by Resource Description Framework (RDF). RDF uses a simple subject-predicate-object statement known as a triple for its basic building block. This provides a much richer exploration of Web and real world resources than the Web of Documents to which we are accustomed.

## LINKED OPEN DATA (LOD) CLOUD



## About the LOD Cloud

The diagram on this page is a visualization of Linked Open Datasets published in the Linked Data format as of April, 2014. The large circle in the center is Dbpedia, the linked data version of Wikipedia. Click on the diagram to learn more about the diagram, licensed and open linked data, statistics about the datasets in the diagram, and the latest version of the LOD Cloud. As of June, 2018, you can view Sub-Clouds by subject area.

Linking Open Data cloud diagram 2014, by Max Schmachtenberg, Christian Bizer, Anja Jentzsch and Richard Cyganiak.

## Contents

### Best Practices, Standards, and Metadata Application Profiles (MAPs)

### 5-Star Open Data Rules

---



### 5-Star Open Data

---

Click on the image of the mug and open the link to access more information.

- [5-Star Open Data](#)  
Click here for an explanation of the costs and benefits of the 5-Star Open Data deployment scheme, and examples.

- [Open Data Certificate](#)

Open Data Institute. Open Data Certificate is a free online tool to assess and recognize the sustainable publication of quality open data. The tool benchmarks data against standards covering legal, practical, technical and social requirements to support the trust in and use of sustainable data. A badge that can be embedded in a website is awarded a data publisher based on answers provided by the publisher to a questionnaire. The Certificate builds on standards such as [opendefinition.org](#), 5\* of Open Data, Sunlight principles, and DCAT.

## Getty Vocabularies Documentation

---

For the Getty Vocabularies, please see the Registries, Portals, and Authorities page under Vocabularies, Ontologies & Frameworks.

## Best Practices and Standards

---

Trust is a major component of the Semantic Web. This requires providing accurate information when publishing a Linked Data instance. The World Wide Web Consortium (W3C), comprised of an international community, develops Web standards and best practices. Additionally, authorities in subject disciplines establish, administer, and maintain standards in their disciplines which adhere to W3C best practices.

This page provides access to information regarding best practices and standards relevant to Semantic Web technology as developed by W3C and other authoritative bodies. For controlled vocabularies, ontologies, etc., please consult the Vocabularies, Ontologies & Frameworks page.

- [ALCTS Standards](#)

Association for Library Collections & Technical Services (ALCTS). The ALCTS Standards is designed to be an aggregator providing a single place to find standards pertinent to the information industry. The guide is organized by topic.

- [Best Practice Recipes for Publishing RDF Vocabularies](#)

Berrueta, Diego and Jon Phipps. (2008, Aug. 28). W3C.

This document describes best practice recipes for publishing vocabularies or ontologies on the Web in RDF Schema or OWL. Each recipe introduces general principles and an example configuration for use with an Apache HTTP server which may be adapted to other environments.

- [Best Practices for Recording Faceted Chronological Data in Bibliographic Records](#)

American Library Association Institutional Repositor Subcommittee on Faceted Vocabularies; Mullin, Casey; Anderson, Karen; Contursi, Lia; McGrath, Kelley; Prager, George; Schiff, Adam. (2020, June 19). This document describes best practices for encoding the date(s) of creation of works and expressions in bibliographic descriptions. The categories of dates, currently serviced by MARC 046 and 388 fields, covered by these practices are: date(s) of creation of individual works; date(s) of creation of the aggregated works in a compilation; date(s) of creation of aggregating works (compilations, anthologies, etc.); and date(s)of creation of expressions.

- [Data on the Web Best Practices](#)

This W3C document provides best practices on a range of topics including data formats, data access, data identification and metadata by providing guidelines on how to represent, describe and make data available in a way that it will be easy to find and to understand. The document provides a series of best practices. A template is used to show the "what", "why" and "how" of each best practice.

- [Generating RDF from Tabular Data on the Web](#)  
W3C. (2015, December 15). This document describes the process of converting tabular data to create RDF subject-predicate-object triples which may be serialized in a concrete RDF syntax such as N-Triples, Turtle, RDFa, JSON-LD, or TriG.
- [Guidelines for Collecting Metadata on Linked Datasets in the datahub.io Data Catalog](#)  
This page explains how data publishers describe datasets they want included in the DataHub (aka LOD Cloud), a registry of open data and content packages maintained by the Open Knowledge Foundation. The page also provides access to a validator that tests whether a data set fulfills the requirements for inclusion in the LOD Cloud.
- [Library of Congress \(LC\) Metadata](#)  
This page provides links to the LC Linked Data Service metadata structure standards including Metadata Authority Description Schema in RDF (MADS/RDF), Simple Knowledge Organization System (SKOS), Web Ontology Language (OWL), Resource Description Framework (RDF), RDF Schema (RDFS), Dublin Core Metadata Initiative Metadata Terms, and SemWeb Vocab Status ontology. There is also an explanation of the relationship between LC authorities and vocabularies and SKOS.
- [Linked Data Platform Best Practices and Guidelines](#)  
This W3C document provides best practices and guidelines for implementing Linked Data Platform [LDP] servers and clients. It also provides links to associated W3C documents.
- [PCC Task Group on URIs in MARC Year One Report](#)  
Bremer, Robert, Folsom, Steven, Frank, Paul, et al. (2016, October 6). This Program for Cooperative Cataloging report discusses the issues associated with setting standards for provisioning URIs in MARC in transitioning from MARC to linked data. Some of the issues include repeatability, pairing, ambiguous relationships, the significance of the ordinal sequence, and identifying a potential field and/or indicator/subfield to record an identifier representing a work.
- [Wikipedia: Authority Control](#)  
Wikipedia. This page describes the editing community's consensus with regard to authority control in Wikipedia articles. It describes how authority control is used in Wikipedia articles to link to corresponding entries in library catalogs of national libraries and other authority files all over the world. The page also provides instruction for using the Wikipedia template to add authority control identifiers to articles.

## Additional Resources about Standards

---

### [Using the W3C Generating RDF from Tabular Data on the Web Recommendation to manage small Wikidata datasets](#)

Baskauf, Steven J. and Baskauf Jessica K. (2021, June 6). This article discusses the W3c recommendation for generating RDF from tabular data.

## Metadata Application Profiles (MAPs)

---

A metadata application profile (MAP) is a set of recorded decisions about a shared application or metadata service, whether it is a datastore, repository, management system, discovery indexing layer, or other, for a given community. MAPs declare what types of entities will be described and how they relate to each other (the model), what controlled vocabularies are used, what fields are required and which fields have a cap on the number of times they can be used, data types for string values, and guiding text/scope notes for consistent use of fields/properties.

A MAP may be a multipart specification, with human-readable and machine-readable aspects, sometimes in a single file, sometimes in multiple files (e.g., a human-readable file that may include input rules, a machine-readable vocabulary, and a validation schema).

The function of a MAP is to clarify the expectations of the metadata being ingested, processed, managed, and exposed by an application or service and document shared community models and standards, and note where implementations may diverge from community standards.

Cornell University Library. (2018, October 23). CUL Metadata Application Profiles. Downloaded January , 2020, from

Library of Congress. (2019, April 30). PCC Task Group on Metadata Application Profiles. Downloaded July 19, 2022 from <https://confluence.cornell.edu/display/mwgweb/CUL+Metadata+Application+Profiles>

- BIBCO Standard Record (BSR) RDA Metadata Application Profile  
Library of Congress, Program for Cooperative Cataloging (PCC). (2017, September 6). The BSR is a model for bibliographic monographic records using a single encoding level (Ldr/17='blank') in a shared database environment, and it follows RDA 0.6.4 in its approach to core. The BSR establishes a baseline set of elements that emphasize access points over descriptive data, while not precluding the use of any data representing a more extensive cataloging treatment. The BSR MAP consists of a combination of RDA Core, RDA Core if, PCC Core, and PCC recommended elements applicable to archival materials, audio recordings, cartographic resources, electronic resources, graphic materials, moving images, notated music, rare materials, and textual monographs. Digital formats, digital reproductions, and authority records are also covered.
- BIBFRAME Profiles: Introduction and Specification  
Library of Congress. (2014, May 5). This document describes how BIBFRAME Profiles are created, maintained and used. It describes an information model and reference serialization to support a means for identifying and describing structural constraints addressing functional requirements, domain models, guidelines on syntax and usage, and possibly data formats.
- PCC Task Group on URIs in MARC Year One Report  
Bremer, Robert, Folsom, Steven, Frank, Paul, et al. (2016, October 6). This Program for Cooperative Cataloging report discusses the issues associated with setting standards for provisioning URIs in MARC in transitioning from MARC to linked data. Some of the issues include repeatability, pairing, ambiguous relationships, the significance of the ordinal sequence, and identifying a potential field and/or indicator/subfield to record an identifier representing a work.
- CONSER Standard Record (CSR) RDA Metadata Application Profile  
Library of Congress, Program for Cooperative Cataloging (PCC). (2020, January 21). The CSR is a model for serial descriptive records using a single encoding level (Ldr/17='blank') in a shared database environment, and it follows RDA 0.6.4 in its approach to the concept of core. The CSR establishes a baseline set of elements that emphasize access points over descriptive data while not precluding the use of any data representing a more extensive cataloging treatment. The CSR consists of a combination of RDA Core, RDA Core if, PCC Core, and PCC Recommended elements applicable to textual serials in various formats. Instructions for rare serials and authority records are included.

- [CUL Metadata Application Profiles](#)  
Cornell University Library Metadata Application Profiles. This page provides an overview and documentation of Cornell University Library's use of metadata application profiles (MAPs). The page offers a definition and explains the role of MAPs in an application or metadata service, and gives examples. A wealth of information regarding documentation for training, MAPS used at CUL, and the CUL metadata ecosystem is provided.
- [DLF AIG Metadata Application Profile Clearinghouse Project](#)  
Digital Library Federation (DLF), Assessment Interest Group (AIG) Metadata Working Group. The mission of this project is to provide a hub and repository for collecting application profiles, mappings, and related specifications that aid or guide descriptive metadata conventions for digital repository collections to be shared with peers in the metadata community. The initial focus is on digital repository descriptive metadata documentation and specifications.
- [Digital Public Library \(DPLA\) Metadata Application Profile](#)  
DPLA MAP Working Group. (2017, December 7). Version 5. This is the technical specification of the DPLA's Metadata Application Profile and provides a list of classes and properties used. Links to other useful documentation include an introduction to the profile, geographic and temporal guidelines, metadata quality guidelines, and rights statements guidelines.
- [Dublin Core Application Profiles \(Guidelines for \)](#)  
This document provides a framework for designing a Dublin Core Application Profile (DCAP), and more generally, a good blueprint for implementing a generic model for metadata records. A DCAP can use any terms that are defined on the basis of RDF, combining terms from multiple namespaces as needed.
- [Dublin Core Collection Description Application Profile](#)  
Dublin Core Collection Description Task Group. (2007, March 9). This document presents full details of the Dublin Core application profile using Dublin Core properties for describing a collection, a catalogue, or an index.
- [IFLA Library Reference Model \(IFLA LRM\)](#)  
International Federation of Library Associations and Institutions (IFLA). (2017, December). IFLA LRM is a high-level conceptual reference model developed within an enhanced entity-relationship modelling framework for bibliographic data. The model aims to make explicit general principles governing the logical structure of bibliographic information, without making presuppositions about how that data might be stored in any particular system or application. Distinctions between data traditionally stored in bibliographic or holdings records and data traditionally stored in name or subject authority records are not made.
- [PCC Task Group on Metadata Application Profiles](#)  
Library of Congress, Program for Cooperative Cataloging (PCC). April 30, 2019. This page outlines the Program for Cooperative Cataloging (PCC)'s Task Group on Metadata Application Profiles charge to help PCC understand issues and practices associated with the management of MAPs and to help develop the expertise needed within PCC to work with MAPs. The charge includes defining MAPs in the PCC context, performing an environmental scan of current work in this space, determining what shareable application profiles means in the PCC context, collaborating with LDRP2 profiles groups, monitoring ongoing LDRP2 PCC Cohort discussions, and recommending actions for a plan to create and maintain profiles that meet stated use cases for application profiles.



### Blogs, Listservs, and Wikis

---

- [BIBFLOW](#)

BIBFLOW is a two-year project of the UC Davis University Library and Zepheira, funded by IMLS. Its official title is “Reinventing Cataloging: Models for the Future of Library Operations.” BIBFLOW’s focus is on developing a roadmap for migrating essential library technical services workflows to a BIBFRAME / LOD ecosystem. This page collects the specific library workflows that BIBFLOW will test by developing systems to allow library staff to perform this work using LOD native tools and data stores. Interested stakeholders are invited to submit comments on the workflows developed and posted on this site. Information from comments will be used to adjust testing as the project progresses.
- [CODE4LIB Wiki](#)

This is the Wiki for library computer programmers and library technologists. It provides information regarding software, conferences, topics, local & regional groups, and interest groups.
- [DBpedia Blog](#)

DBpedia is an open, free, and comprehensive global knowledge base which is continuously extended and improved by putting into effect a quality-controlled and reliable fact extraction from Wikipedia and Wikidata. This blog provides information regarding DBpedia, tools, events, dataset releases, the the DBpedia ontology, and more.
- [Dublin Core Metadata Initiative Wiki](#)

This MediaWiki for the Dublin Core Metadata Initiative (DMCI) provides information on DCMI's activities regarding work on architecture and modeling, discussions and collaborative work in DCMI Communities and DCMI Task Groups, annual conferences and workshops, standards liaison, and educational efforts to promote widespread acceptance of metadata standards and practices. Access the DCMI Handbook and LD4PE Linked Data Exploratorium.
- [FRBR Open Comments](#)

This blog encourages transparency and invites comments regarding the continued development of the international library entity relationship model, the Functional Requirements of Bibliographic Records (FRBR) and the FRBR-Library Reference Model (FRBR\_LRM), a consolidation of the FRBR, FRAD and FRSAD conceptual models. Access an Executive Summary, and read or contribute to the General Comments or other areas of interest such as User tasks, Entities, User population considered, Entity-Relationship Diagrams, Modeling of Aggregates, and more.
- [Hanging Together: The OCLC Research Blog](#)

Hanging Together is OCLC's research blog. It provides information about the types of projects and issues which OCLC is researching and with whom it is partnering. The blog covers a wide range of topics including Architecture and Standards, Digitization, Identifiers, Infrastructure, Linked Data, Metadata, Modeling New Services, and more.
- [Schema Bib Extend Community Group](#)

This is the main Wiki page for the Schema Bib Extend Community Group, a W3C group formed to discuss and prepare proposal(s) for extending Schema.org schemas for the improved representation of bibliographic information markup and sharing. The Wiki provides links to the following topics: Recipes and Guidelines for those looking to adopt Schema.org for bibliographic data; Areas for Discussion; Use Cases; Scope; Object Types; Vocabulary Proposals; and Example Library.

- [Schema blog](#)  
This is the official schema.org blog.

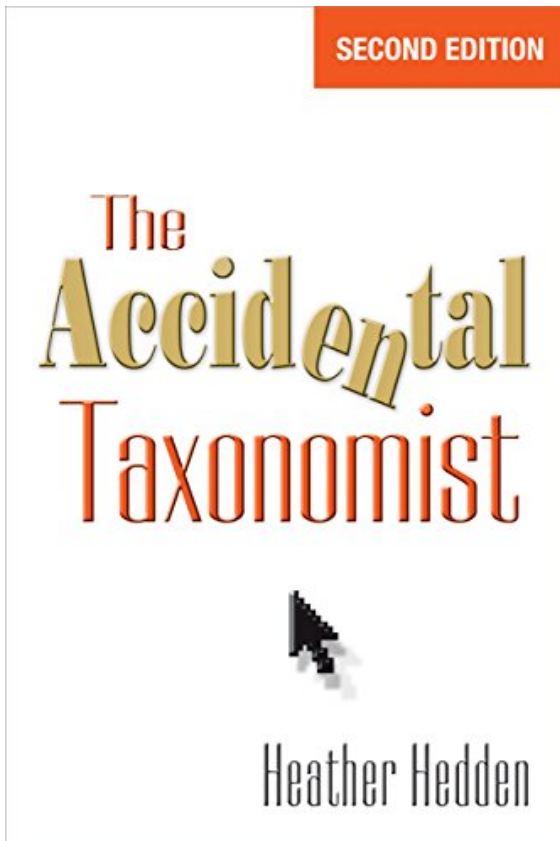
## Books

### Books

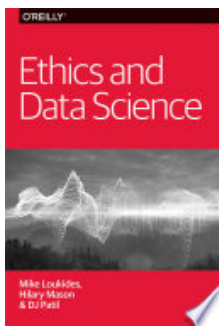
---

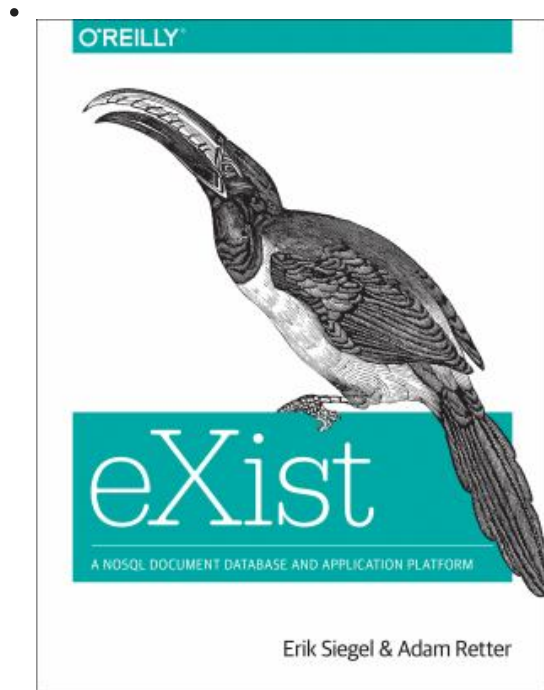
Below is a list of books which provide a good introduction to the Semantic Web. Items whose titles are highlighted in blue link either to the UCLA Library record for that title if the tile is held by the library, or to an online copy if available. Use the Safari Books Online link to search for additional resources.

- [The Accidental Taxonomist by Heather Hedden](#)  
ISBN: 9781573875288  
Date: 2016-06-13



- [Ethics and Data Science by by Mike Loukides, Hilary Mason, DJ Patil](#)  
ISBN: 9781492043881  
Date: July, 2018  
This book examines practical ways for making ethical data standards part of your work.



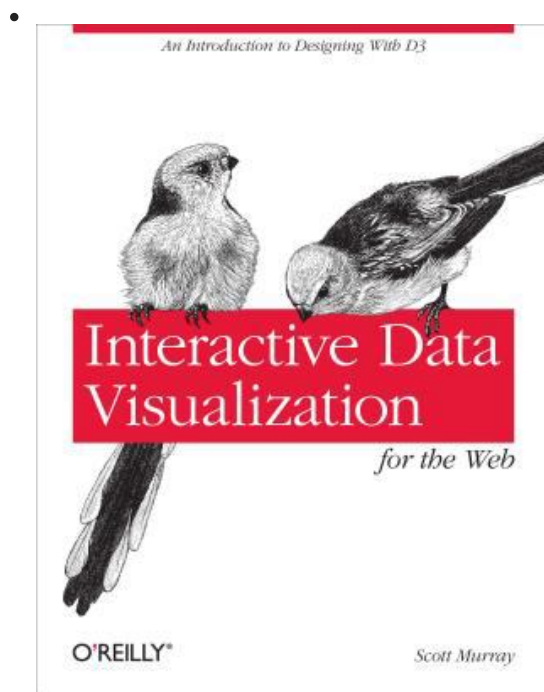


[eXist](#) by Erik Siegel; Adam Retter

ISBN: 9781449337100

Date: 2014-12-25

This is a hands-on guide for eXist, an open source NoSQL database and application development platform built around XML and XQuery technologies.



[Interactive Data Visualization for the Web, 2nd edition](#) by

[Scott Murray](#)

ISBN: 9781491921289

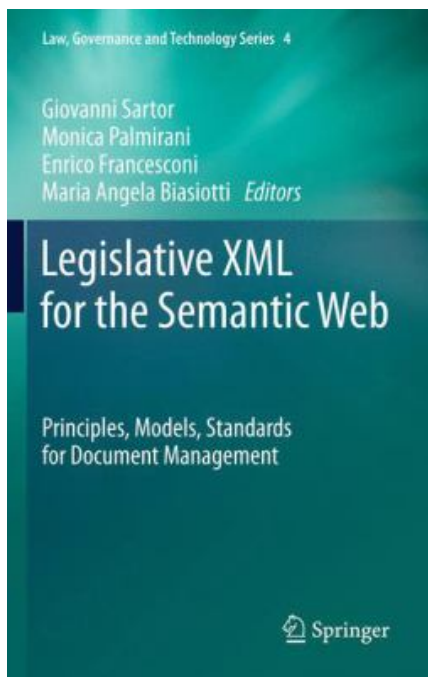
Create and publish interactive data visualization projects on the Web in a web browser using HTML, CSS, JavaScript, and SVG basics to dynamically generate web page elements from your data.



[Learning SPARQL by Bob DuCharme](#)

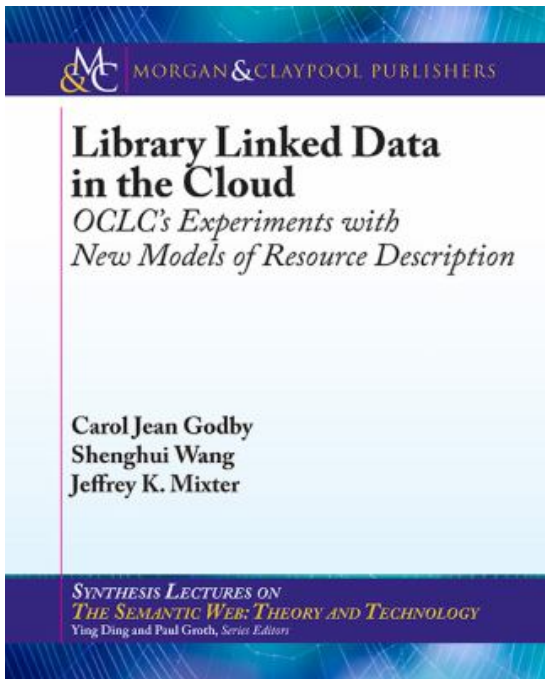
ISBN: 1449371434

The book demonstrates how to use SPARQL 1.1 with a variety of tools to retrieve, manipulate, and federate data from the public web as well as from private sources.



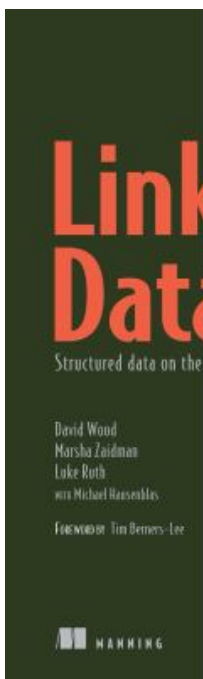
[Legislative XML for the Semantic Web by Giovanni Sartor \(Editor\); Maria Angela Biasiotti \(Editor\); Enrico Francesconi \(Editor\); Monica Palmirani \(Editor\)](#)

ISBN: 9789400718876



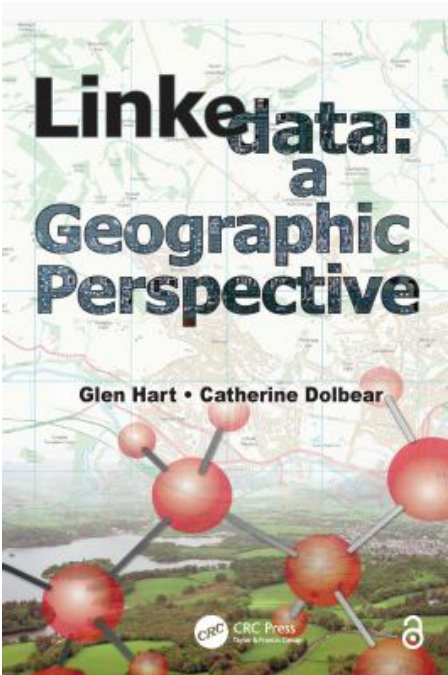
Library linked data in the cloud by Carol Jean Godby, Shenghui Wang, Jeffrey K. Mixer  
ISBN: 1627052194

Full title: Library linked data in the cloud : OCLC's experiments with new models of resource description.  
Follow the link for access to an open access PDF version.



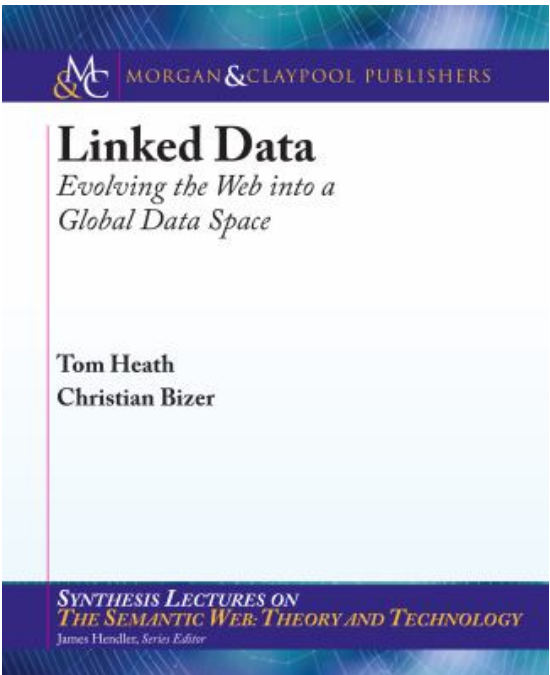
Linked Data by David Wood; Marsha Zaidman; Luke Ruth; Michael Hausenblas  
ISBN: 1617290394

This book presents practical techniques for using Linked Data on the Web using familiar tools like JavaScript and Python and introduces foundational concepts such as HTTP URIs, the Resource Description Framework (RDF), and the SPARQL query language.

- 


[Linked Data: A geographic perspective by Glen Hart; Catherine Dolbear](#)  
 ISBN: 1439869952

Although written for a geographic oriented audience, this book provides one of the best introductions to basic Linked Data concepts.

- 

[Linked Data: Evolving the Web into a Global Data Space by Tom Heath; Christian Bizer](#)  
 ISBN: 9781608454310

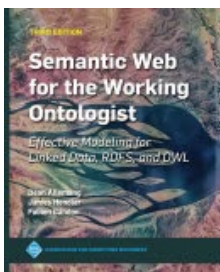
This books provides guidance and best practices regarding architectural approaches to publishing Linked Data on the Web.

- 

[Linked Data for Libraries, Archives and Museums: How to Clean, Link and Publish your Metadata by Seth van Hooland, Ruben Verborgh](#)  
 ISBN: 1783303018



- [Semantic Web for the Working Ontologist by Dean Allemang; James Hendler](#)  
ISBN: 1450376150



This book provides an introduction to Resource Description Framework Schema (RDFS), Web Ontology Language (OWL), SPARQL for accessing RDF data, Simple Knowledge Organization System (SKOS), and other semantic web concepts and modeling.

- [The Taxobook, Part 2 by Marjorie Hlava](#)  
ISBN: 3031022882



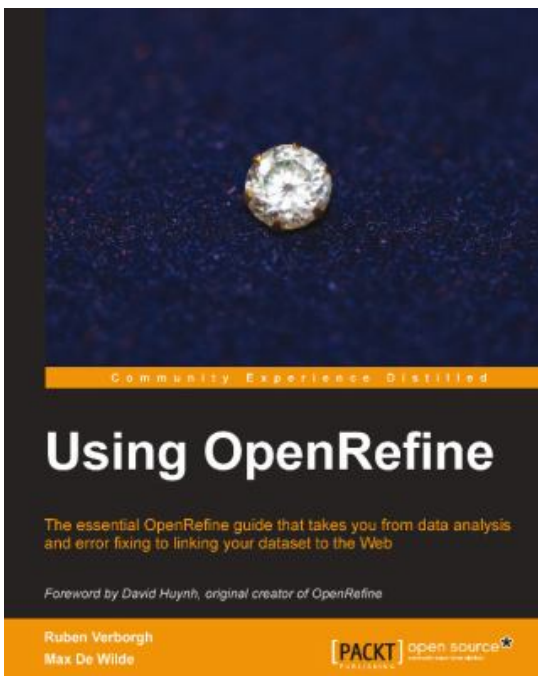
Part 2 of 3. This book outlines the basic principles of creation and maintenance of taxonomies and thesauri. It also provides step by step instructions for building a taxonomy or thesaurus and discusses the various ways to get started on a taxonomy construction project.

- [The Taxobook by Marjorie M. K. Hlava](#)  
ISBN: 3031022904



This book is Part 3 of a three-part series on taxonomies, and covers putting your taxonomy into use in as many ways as possible to maximize retrieval for your users.

- [Using OpenRefine by Ruben Verborgh; Max De Wilde](#)  
Call Number: QA76.9.D343



ISBN: 9781783289080

Date: 2013-09-10

The book is styled on a Cookbook, containing recipes - combined with free datasets - which teaches readers how to use OpenRefine to clean and handle large amounts of data. Datasets to use with the learning "recipes" are provided.

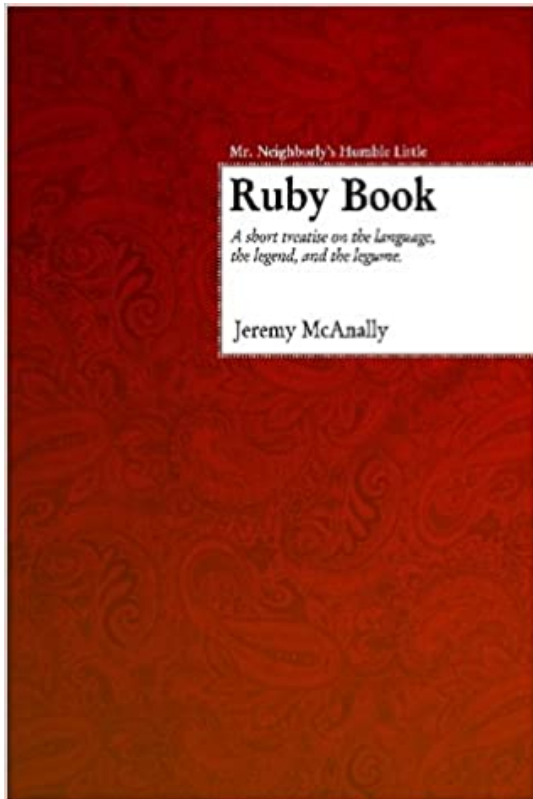
- 



[Linked Data Patterns: A pattern catalogue for modelling, publishing, and consuming Linked Data](#) by Leigh Dodds and Ian Davis

Organized as a pattern catalogue, this book serves as a ready reference for the beginner and the experienced semantic publisher. It is a good companion to the book, "Semantic Web for the Working Ontologist."

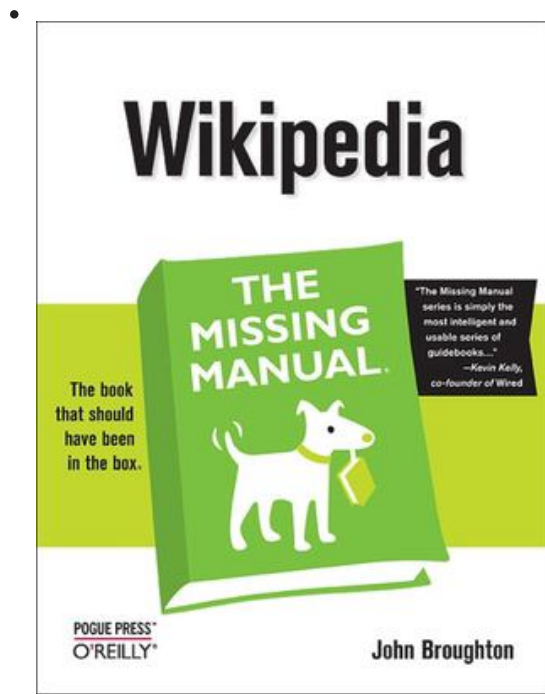
- 



[Mr. Neighborly's Humble Little Ruby Book](#)

Ruby is an open-source, multi-paradigm, interpreted programming language. If you don't know what any of that means, this is the book to introduce you to programming Ruby.





[Wikipedia: The Missing Manual by John Broughton](#)  
 ISBN: 0596515162

Learn how to contribute to Wikipedia, the user-generated online reference encyclopedia. Wikipedia: The Missing Manual gives you practical advice on creating articles and collaborating with fellow editors, improving existing articles, and working with the Wikipedia community to review new articles, mediate disputes, and maintain the site.



[Linked Open Data: The Essentials – The Climate Knowledge Brokering Edition by Bauer, Florian; Kaltenböck, Martin](#)  
 Date: 2016

## Datasets

## Datasets

This page provides a short list of datasets and data portals. To explore the global network of datasets connected on the Web, click on the Linked Open Data Cloud on the home page.

- [DataCite](#)  
 DataCite is a global non-profit organization that provides persistent identifiers (DOIs) for research data and other research outputs. Use it to locate, identify, and cite research data. DataCite provides several services including a global registry of research data repositories from a diverse range of academic disciplines and information about them (re3data.org), a citation formatter, content negotiation, a Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) service, and more.

- [Data.gov](#)  
This page provides access to the datasets in the United States open government data catalog. Data are provided by hundreds of organizations and Federal agencies. It provides an online repository of policies, tools, case studies, and other resources to support data governance, management, exchange, and use throughout the federal government.
- [Data Hub - Linking Open Data Cloud](#)  
This Data Hub group catalogs data sets that are available on the Web as Linked Data and contain data links pointing at other Linked Data sets. A search option for the datasets is available. The descriptions of the data sets in this group are used to generate the Linking Open Data Cloud diagram at regular intervals. The descriptions are also used generate the statistics provided in the State of the LOD Cloud document. The descriptions are also used generate the statistics provided in the State of the LOD Cloud document.
- [Data Portals](#)  
DataPortals.org is a comprehensive list of government and NGO open data portals across the world. It is curated by a group of leading open data experts from around the world, including representatives from local, regional and national governments, international organizations such as the World Bank, and numerous NGOs.
- [DBpedia](#)  
DBpedia is a crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the Web. DBpedia provides the ability for sophisticated queries against Wikipedia, and to link the different data sets on the Web to Wikipedia data.
- [EPSG Geodetic Parameter Dataset](#)  
Geodesy Subcommittee of the International Association of Oil & Gas Producers (IOGP). The EPSG Geodetic Parameter Dataset is a structured dataset of Coordinate Reference Systems and Coordinate Transformations. It can be accessed through an online registry or downloaded as zip files. Geographic coverage is worldwide, but it does not record all possible geodetic parameters in use around the world. The dataset is maintained by the IOGP's Geomatics Committee.
- [Europeana](#)  
Europeana provides access to European cultural heritage material from institutions across Europe. Discover artworks, books, music, and videos on art, newspapers, archaeology, fashion, science, sport, and much more.
- [GOKb](#)  
GOKb (Global Open Knowledge base) is an an open data repository to describe electronic journals and books, publisher packages, and platforms for use in a library environment. It includes tracking changes over time, including publisher take-overs and bibliographic changes.
- [Linked Open Data Cloud](#)  
lod-cloud.net. This is the home of the LOD Cloud diagram. It is a dataset of datasets published in Linked Data format contained in the LOD Cloud. Datasets contained in the Cloud should follow the Linked Data principles listed on the site's About page. Subject areas have been broken into Subclouds for easier use.
- [List of online music databases](#)  
Wikipedia. (2021, April 19). This page lists music domain datasets covering sheet music, reviews, artists, labels, a heavy metal encyclopedia, audio samples, a database of Arabic and Middle Eastern music artists, tracks, and albums, biographies and discographies, audio based music recognition and provision of song lyrics, and more.

- [Resources.data.gov](https://resources.data.gov)  
This repository of Federal enterprise data resources provides links to policies, tools, case studies, and other resources to support Federal government data governance, management, exchange, and use.
- [WordNet](https://wordnet.princeton.edu/)  
WordNet® is a lexical database of English useful for computational linguistics and natural language processing. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets). Synsets are interlinked by means of conceptual-semantic and lexical relations. The resulting network of meaningfully related words and concepts can be navigated with the browser. The dataset is available for downloading. Unfortunately, due to staffing, updates have been suspended.

## Instructional Resources

### Instructional Resources

---

There are many resources available to help you learn about the Semantic Web and Linked Data. This page provides access to a few instructional resources on topics relating to Linked Data in a variety of formats. See the SPARQL page for SPARQL related instructional resources.

- [BIBFRAME Manual](#)  
Library of Congress. (2019). This is the Library of Congress training manual for the BIBFRAME Editor and BIBFRAME Database.
- [BIBFRAME Training at the Library of Congress](#)  
The Library of Congress is providing training for participants in the BibFrame Pilot which is testing bibliographic description in multiple formats and in multiple languages. This website provides access to the three training modules: 1) Introduction to the Semantic Web and Linked Data; 2) Introduction to the BibFrame Tools; and 3) Using the BibFrame Editor. There is a PowerPoint presentation and quiz for each module, and some modules have additional resources.
- [Catalogers Learning Workshop \(CLW\)](#)  
Library of Congress. This page links to Library of Congress training materials for topics such as Library of Congress Subject Headings, RDA: Resource Description & Access; BIBFRAME training at the Library of Congress; BIBFRAME Webcasts and Presentations; and other training resources.
- [Competency Index for Linked Data \(CI\)](#)  
LD4PE. The Competency Index for Linked Data (CI) is an initiative of Exploring Linked Data, a Linked Data for Professional Educators (LD4PE) project. The web site supports the structured discovery of learning resources for Linked Data available online by open educational resource (OER) and commercial providers. The site indexes learning resources within a framework according to specific competencies, skills, and knowledge they address. Tutorials are available for such topics as Fundamental of Resource Description Framework (RDF), Fundamentals of Linked Data, RDF Vocabularies and Application Profiles, Creating and Transforming Linked Data, Interacting with RDF data, and Creating Linked Data applications. LD4PE is administered under the jurisdiction of the DCMI Education & Outreach Committee and is funded by the Institute of Museum and Library Services (IMLS).

- [Free Your Metadata](#)  
This site, geared for libraries, archives, and museums, enables the matching of metadata with controlled vocabularies connected to the Linked Data cloud and the enriching of unstructured description fields using the named entity extraction tool OpenRefine extension. Learn how to check for errors and correct them, and publish metadata in a sustainable way. The site also provides information on relevant publications.
- [The language of languages](#)  
Might, Matt. This article provides a brief explanation of grammars and common notations for grammars, such as Backus-Naur Form (BNF), Extended Backus-Naur Form (EBNF) and regular extensions to BNF. Grammars determine the structure of programming languages, protocol specifications, query languages, file formats, pattern languages, memory layouts, formal languages, config files, mark-up languages, formatting languages, and meta-languages. The Extended Backus-Naur Form notation is used to describe the essential BIBFRAME Profile syntax elements.
- [Linked Data: Evolving the Web into a Global Data Space](#)  
Tom Heath, Tom and Bizer, Christian. (2011). (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool. This overview of Linked Data principles and the Web of Data discusses patterns for publishing Linked Data and describes deployed Linked Data applications and their architecture. This book supersedes the publication, "How to Publish Linked Data on the Web," by Chris Bizer, Richard Cyganiak, and Tom Heath.
- [Linked Data Tools](#)  
This site has been created by professional developers to help the web community transition into Web 3.0, or the Semantic Web. The site provides tools and tutorials for learning how to begin using the semantic web.
- [MarcEdit and OpenRefine](#)  
Reese, Terry. (2016, January 16). This page describes how to export a MARC file for use in OpenRefine.
- [MARCEdit You Tube Videos](#)  
This page lists over 90 videos produced by Terry Reese providing instructions for using MARCEdit. Topics include "MarcEdit 101: I have a MARC record, now what?," "Installing MarcEdit natively on a Mac operating system," "Extract and Edit Subsets of Records in MarcEdit," "MarcEdit Task Automation Tool," and "MarcEdit RDA Helper."
- [NCompass Live: Metadata Manipulations: Using MarcEdit and OpenRefine](#)  
Nebraska Library Commission. (2015, June 24). This tutorial provides instruction for using OpenRefine and MARCEdit.
- [NCompass Live: Metadata Manipulations: Using Marc Edit And Open Refine To Enhance Technical Services Workflows](#)  
Nebraska Library Commission. (2015, June 24). This video shows how to use MARCEdit and OpenRefine to edit your catalog records more efficiently, transform your library data from one format to another, and detect misspellings and other inaccuracies in your metadata.
- [Ontogenesis](#)  
Lord, Phillip. (2012). This is an archived Knowledge Blog which provides access to descriptive, tutorial, and explanatory material about building, using, and maintaining ontologies, as well as the social processes and technology that support this. There are links to articles, many peer reviewed, and tutorials regarding a range of topics of interest for developers and users of ontologies.

- [Ontology Development 101: A Guide to Creating Your First Ontology](#)  
Noy, Natalya F. and McGuinness, Deborah L. Stanford University. This guide discusses the reasons for developing an ontology and the methodology for creating an ontology based on declarative knowledge representation systems.
- [OpenRefine Wiki External Resources](#)  
This page lists tutorials and resources developed outside the OpenRefine wiki covering a wide range of topics and use cases, including general instruction, data clean up, geospatial metadata, spreadsheet transformations, and much more.
- [Programming Historian](#)  
Crymble, Adam, Fred Gibbs, Allison Hegel, Caleb McDaniel, Ian Milligan, Evan Taparata, and Jeri Wieringa, eds. (2016). The Programming Historian. 2nd ed. This blog provides peer-reviewed tutorials geared towards helping humanists learn a wide range of digital tools, techniques, and workflows to facilitate their research. Several of the tutorials are related to linked data. Other tutorials may be of interest to those generating or consuming data.
- [RDFa with schema.org codelab: overview](#)  
Scott, Dan. (2014, Dec.1). This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.  
Using detailed instructions and examples, this page walks through the process of using schema.org to enhance library web pages so that they contain structured data using the schema.org vocabulary and RDFa attributes.
- [Semantic Web Data at the University of Washington Libraries](#)  
Cataloging and Metadata Services, University of Washington. This webpage links to a wide range of useful resources and guidelines for working with Linked Data in a University setting. The project was developed by the Institute of Museum and Library Services.
- [What Can We Do About Our Legacy Data?](#)  
Hillmann, Diane. (2015). This is Diane Hillmann's presentation given at the 2015 American Library Association Conference raising questions about moving library data onto the Semantic Web. Posted to SlideShare on June 29, 2015.
- [XPath Tutorial](#)  
This W3schools page provides an introductory tutorial for XPath, a language for finding information in an XML document.

## Journals, Articles, and Papers

## **Journals, Articles, and Papers**

---

- Bulletin of the Association for Information Science and Technology  
Association for Information Science and Technology. Silver Spring, Maryland. [2013-]

- **Cataloging & Classification Quarterly**  
Haworth Press. Binghamton, NY. (1981)- ISSN: 1544-4554. The journal covers the full spectrum of creation, content, management, use, and usability of bibliographic records, including the principles, functions, and techniques of descriptive cataloging. The range of methods of subject analysis and classification, provision of access for all formats of materials, and policies, planning, and issues connected to the effective use of bibliographic data in modern society are also focuses of this journal.
- **The Code{4}Lib Journal**  
Code{4}Lib Journal, Chapel Hill, N.C.. (2007-). ISSN: 1940-5758The focus of this journal is to provide the library community with information regarding technology tools for managing information in libraries.
- **International Journal of Web & Semantic Technology**  
Academy & Industry Research Collaboration Center (AIRCC). (2010 - .)  
ISSN: 0975-9026; EISSN: 0975-9026. This journal focuses on theory, methodology, and applications of web and semantic technology.
- **Journal of library metadata**  
Haworth Press. New York, NY. (2008 - ). SSN : 1937-5034; ISSN : 1938-6389. The metadata that describes library resources is becoming more critical for digital resource management and discovery. This journal covers application profiles, best practices, controlled vocabularies, cross walking of metadata and interoperability, digital libraries and metadata, federated repositories and searching, folksonomies, individual metadata schemes, institutional repository metadata, metadata content standards, resource description framework, SKOS, topic maps, and more.
- **Journal of the Association for Information Science and Technology**  
Association for Information Science and Technology. Wiley Blackwell. Hoboken, NJ. (2014). This journal publishes original research that focuses on the production, discovery, recording, storage, representation, retrieval, presentation, manipulation, dissemination, use, and evaluation of information and on the tools and techniques associated with these processes.
- **Library Technology Reports**  
American library Association, Chicago, Ill. (2009 - ). Library Technology Reports focuses on the application of technology to library services, including evaluative descriptions of specific products or product classes and covers emerging technology. The journal is sunsetting December, 2022 and will be available for single-issue sales only.
- **Web Semantics : Science, Services and Agents on the World Wide Web**  
Elsevier Science. Amsterdam; New York. (2004)- ISSN: 1873-7749; ISSN : 1570-8268. This journal covers all aspects of Semantic Web development including topics such as knowledge technologies, ontology, agents, databases and the semantic grid. It also focuses on disciplines such as information retrieval, language technology, human-computer interaction and knowledge discovery.

## Articles and Papers

---

- **Addressing the Challenges with Organizational Identifiers and ISNI**  
Smith-Yoshimura, Karen, Gatenby, Janifer, Agnew,Grace, Brown,Christopher, Byrne, Kate, Carruthers,Matt, Fletcher, Peter, Hearn, Stephen, Li, Xiaoli, Muilwijk, Marina, Naun, Chew Chiat, Riemer, John, Sadler, Roderick, Wang, Jing, Wiley, Glen, and Willey, Kayla. (2016). Dublin, Ohio: OCLC Research. This paper discusses a model for using unique identifiers that are resolvable globally over networks via a specific protocol to provide the means to find and identify an organization accurately and to define the relationships among its sub-units and with other organizations.

- [A Division of Labor: The Role of Schema.org in a Semantic Web Model of Library Resources](#)  
Godby, Carol Jean. (2017). This article describes experiments with Schema.org conducted by OCLC as a foundation for a linked data model for library resources, and why Schema.org was the vocabulary considered in designing the next generation standards for library data.
- [Creating Organization Name Authority within an Electronic Resources Management System](#)  
Blake, K., & Samples, J. (2009) *Library Resources & Technical Services*, 53(2), 94-107. To access the linked data project associated with this article, click on Organization Name Linked Data on our Use Cases Page.
- [Creating Value with Identifiers in an Open Data World](#)  
Open Data Institute and Thomson Reuters. (2014) *Creating Value with Identifiers in an Open Data World*. Retrieved from <http://thomsonreuters.com/site/data-identifiers>. This joint effort between Thomson Reuters and the Open Data Institute serves as a guide for how identifiers can create value by empowering linked data for publishing and discovery.
- [The Global Open Knowledgebase \(GOKb\): open linked data supporting electronic resources management and scholarly communication](#)  
Antelman, Kristin and Wilson, Kristen. (2015). DOI: <http://doi.org/10.1629/uksg.217>. CC BY 3.0 License. Kristen Wilson Global Open Knowledgebase is an open data repository of information related to e-resources as they are acquired and managed in a library environment. This article describes how the GOKb model was developed to track this information.
- [Hello BIBFRAME2.0: Changes from 1.0 and Possible Directions for the Future](#)  
Kroeger, Angela. J. (2016, October 20). *Criss Library Faculty Proceedings & Presentations*. 65. This presentation introduces the basics and history of the BIBFRAME model, and its relationship to RDF, FRBR, and RDA. It covers core classes, editors, mixing metadata, holdings, approaches, PREMIS, changes from BIBFRAME1.0, and more.
- [Introducing the FRBR Library Reference Model](#)  
Riva, Pat, and Žumer, Maja. (2015). This paper serves as an introduction to the FRBR Library Reference Model which consolidates the FRBR, FRAD, and FRSAD models for bibliographic data, authority data, and subject authority data so that the model's definitions can be readily transferred to the IFLA FRBR namespace for use with linked open data applications.
- [Linked Data in Libraries: A Case Study of Harvesting and Sharing Bibliographic Metadata with BIBFRAME](#)  
Tharani, Karim. (2015). In *"Information Technology and Libraries"*, 34(1). This paper illustrates and evaluates the Bibliographic Framework (BIBFRAME) as a means for harvesting and sharing bibliographic metadata over the web for libraries. With BIBFRAME disparate library metadata sources such as catalogs and digital collections can be harvested and integrated over the web.
- [LTS and Linked Data: a position paper](#)  
Naun, Chew Chiat, Kovari, Jason, and Folsom, Steven. (2015, Dec. 16). Prepared for Cornell University Library Technical Services (LTS), this paper explores reasons for adopting linked data techniques for describing and managing library collections, and seeks to articulate a specific role for Library Technical Services within this linked data environment.
- [Making Ontology Relationships Explicit in a Ontology Network](#)  
Díaz, Alicia, Motz, Regina, and Rohrer, Edelweis. (2011). This paper formally defines the different relationships among networked ontologies and shows how they can be modeled as an ontology network in a case study of the health domain.

- RDA vocabularies for a twenty-first-century data environment  
Coyle, Karen. (2010). Library technology reports, v. 46, no. 2, p.5-39. Contents include Library Data in the Web World, Metadata Models of the World Wide Web, FRBR, the Domain Model, and RDA in RDF.
- The Relationship between BIBFRAME and OCLC's Linked-Data Model of Bibliographic Description: A Working Paper  
Godby, Carol Jean. (2013, June). Dublin, Ohio: OCLC Research. This paper describes a proposed alignment between BIBFRAME and an OCLC model using Schema Bib Extend extensions to enhance Schema.org for use with the description of library resources.
- Sharing Research Data and Intellectual Property Law: A Primer  
Carroll, Michael W. (2015) PLoS Biol 13(8): e1002235. doi:10.1371/journal.pbio.1002235.  
This article explains how to work through the general intellectual property and contractual issues for all research data.
- Towards Identity in Linked Data  
McCusker, James P. and McGuinness, Deborah L. Rensselaer Polytechnic Institute.  
This paper poses problems with and solutions for using owl:sameAs for linking datasets when dealing with provenance, context, and imperfect representations in Linked Data. The paper uses examples of merging provenance in biomedical applications.
- Understanding Metadata  
Riley, Jenn. National Information Standards Organization (NISO). This primer serves as a guidance for using data and covers developments in metadata, new tools, best practices, and available resources.
- Web-Scale Querying through Linked Data Fragments  
Verborgh, Ruben, Vander Sande, Miel, Colpaert, Pieter, Coppens, Sam, Mannens, Erik, Van de Walle, Rik. (2014). This paper explains the core concepts behind Linked Data Fragments, a method that allows efficient linked data query execution from servers to clients through a lightweight partitioning strategy.
- When owl:sameAs Isn't the Same: An Analysis of Identity in Linked Data  
Halpin, Harry, Hayes, Patrick J., McCusker, James P., McGuinness, Deborah L., and Thompson, Henry S. (2010). Patel-Schneider, P. F. et al. (Eds.): ISWC 2010, Part I, LNCS 6496, pp. 305–320, Springer-Verlag Berlin Heidelberg. This document discusses how owl:sameAs is being used and misused on the Web of data, particularly with regards to interactions with inference. The authors describe how referentially opaque contexts that do not allow inference exist, and outline some varieties of referentially-opaque alternatives to owl:sameAs.

## **Semantic Web Services**

## **Semantic Web Services**

---

This page lists Semantic Web services which are of interest to information specialists, libraries, museums, and cultural organizations.

- Library.Link Network  
Library.Link Network is a service which transforms data from library resources into searchable resources on the Web using Linked Data.



- [Library of Congress Linked Data Service](#)  
This is the portal for all of the Library of Congress' Linked Data Vocabularies and Authorities, including without limitation, LC Subject Headings, Name Authority File, MARC Relators, LC Classification, LC Children's Subject Headings, LC Genre/Form Terms, ISO Languages, Cultural Organizations, Content Types, to name a few.
- [Share-VDE](#)  
Share-VDE (SVDE) is a discovery interface offering an intuitive delivery service of wide-ranging and detailed search results to library patrons. Library catalogues of participating institutions are converted from MARC to Resource Description Framework (RDF) using the BIBFRAME vocabulary and other ontologies to form clusters of entities. The network of resources created is published as linked data. A common knowledge base of clusters is compiled in a Cluster Knowledge Base named Sapiaientia. Participating libraries handle their own data as independently as possible and receive their original records converted into linked data. The SVDE infrastructure is built on the LOD Platform.
- [VIAF: The Virtual International Authority File](#)  
VIAF links and matches multiple name authority files from global resources into a single OCLC-hosted name authority service increasing the utility of library authority files and making them available on the Web.
- [WorldCat Entities](#)  
OCLC. (2022). This OCLC service provides the ability to search WorldCat Entities for persons and works. Browse through different languages and explore the way each entity links to other external vocabularies and authority.

## **Semantic Web Tools**

### **Semantic Web Tools**

---

Semantic Web technology uses an array of tools. This page lists conversion tools, data management tools, glossaries, ontology & vocabulary building platforms, Semantic Web browsers, validators, XML editors, and XPath tools.

#### W3C Semantic Web Tools

This Wiki lists an array of tools for developing Semantic Web applications compiled by the W3C, including development environments, editors, libraries or modules for various programming languages, specialized browsers, and more.

### **Assessment Tools**

---

- [DLF AIG MWG Metadata Assessment Toolkit](#)  
The Digital Library Federation (DLF) Assessment Interest Group (AIG) Metadata Working Group (MWG) aka DLF Metadata Assessment Working Group. The toolkit is a great resource for assessment information and tools and covers a review of the literature, tools, and organizations concerning metadata assessment, quality, and best practices. The site provides a list of metadata assessment tools, and a collection of application profiles, mappings, code and best practices provided by several institutions.

- [LODQuator](#)  
LODQuator is a data portal built on the Luzzu Quality Assessment Framework for ranking and filtering Linked Open Data Cloud datasets. It provides the ability to search datasets based on their quality using over a dozen metrics which are listed on the site.
- [Luzzu](#)  
Enterprise Information Systems (EIS) at Fraunhofer Institute for Intelligent Analysis and Information Systems (IAIS), University of Bonn. Luzzu is a quality assessment framework for Linked Data Open datasets based on the Dataset Quality Ontology (daQ). It assesses Linked Data quality using user-provided domain specific quality metrics in a scalable manner, provides query enabled quality metadata on assessed datasets, and assembles detailed quality reports on assessed datasets.
- [Open Data Certificate](#)  
Open Data Institute. Open Data Certificate is a free online tool to assess and recognize the sustainable publication of quality open data. The tool benchmarks data against standards covering legal, practical, technical and social requirements to support the trust in and use of sustainable data. A badge that can be embedded in a website is awarded a data publisher based on answers provided by the publisher to a questionnaire. The Certificate builds on standards such as opendefinition.org, 5\* of Open Data, Sunlight principles, and DCAT.

## Authority Tools

---

### [Authority toolkit: create and modify authority records](#)

Strawn, Gary L. (2016, June 30). Northwestern University. Evanston, IL USA. This document describes how the Authority Toolkit can be used to create a new authority record from an access field in a bibliographic record. Use the tool to help you enhance the preliminary authority record, enhance an existing authority record, or extract one identity from an undifferentiated personal name authority record and then enhance the preliminary authority record for the extracted identity. The tool can be used to extract information from sources such as VIAF, Wikidata, Wikipedia, and the CERL thesaurus into authority records.

## BIBFRAME Tools

---

- [BIBFRAME Comparison Tool](#)  
This tool provides for the side-by-side conversion of MARCXML records from the Library of Congress database to BIBFRAME2 using a LCCN or record number. Records can be serialized in Turtle or RDF XML.
- [Bibliographic Framework Initiative](#)  
Library of Congress. The Bibliographic Framework Initiative is the replacement for MARC developed by the Library of Congress and is investigating all aspects of bibliographic description, data creation, and data exchange. More broadly the initiative includes accommodating different content models and cataloging rules, exploring new methods of data entry, and evaluating current exchange protocols. This page provides access to the BIBFRAME 2.0 model, vocabulary, extension list view, and MARC 21 to BIBFRAME conversion tools. The BIBFRAME Implementation Register can be accessed here.
- [marc2bibframe2](#)  
This tool, available on GitHub, uses an XSLT 1.0 application to convert MARCXML to RDF/XML, using the BIBFRAME 2.0 and MADSRDF ontologies. Information regarding integration of the application with Metaproxy is also available.

- [MARC 21 to BIBFRAME 2.0 Conversion Specifications](#)

These specifications were developed to support a pilot in the use of BIBFRAME 2.0 at the Library of Congress. They specify the conversion of MARC Bibliographic records to BIBFRAME Work, Instance and Item descriptions, and MARC Authority records for titles and name/titles to BIBFRAME Work descriptions. The specifications were written from the perspective of MARC so that each element in MARC would at least be considered, even if not converted. The specifications are presented in MS Excel files with explanatory specifications in MS Word.

- [Sinopia](#)

Sinopia is an implementation of the Library of Congress BIBFRAME Editor and Profile Editor.

## Conversion Tools

---

- [Freeformatter JSON to XML Converter](#)

This tool converts a JSON file into an XML file. The converter uses rules to make allowances for XML using different item types that do not have an equivalent JSON representation.

- [Freeformatter XML to JSON Converter](#)

This tool converts an XML file into a JSON file. The converter uses rules to make allowances for XML using different item types that do not have an equivalent JSON representation.

- [OxGarage](#)

OxGarage is a web, RESTful conversion service developed by the University of Oxford IT Services. The majority of transformations use the Text Encoding Initiative (TEI) format as a pivot format, and many other formats are supported, including TEI to Word and Word to TEI. Give the page a moment to load. Choose a format from a menu of Documents, Presentations, or Spreadsheets to convert to a format from a list provided for each menu option.

- [Pandoc](#)

Pandoc converts documents in markdown, reStructuredText, textile, HTML, DocBook, LaTeX, MediaWiki markup, TWiki markup, OPML, Emacs Org-Mode, Txt2Tags, Microsoft Word docx, LibreOffice ODT, EPUB, or Haddock markup to HTML formats, word processor formats, Ebooks, documentation formats, page layout formats, outline formats, TeX formats, PDF, lightweight markup formats, and custom formats.

- [SearchFAST](#)

OCLC Research. SearchFast is a suite of tools for working with FAST headings. The tools include a converter to convert Library of Congress Subject Headings to FAST headings, searchFast, a search interface for the FAST database, and mapFast, a Google Maps mashup to provide map based access to bibliographic records using FAST geographic and event authorities. Other tools in the suite include FAST Linked Data, authorities formatted using schema.org and SKOS (Simple Knowledge Organization System) that are linked to LCSH and other authorities such as VIAF, Wikipedia, and GeoNames, and assignFast, a web service that automates manual selection of FAST subjects.

## Data Management Tools

---

- [CKAN](#)  
CKAN (Comprehensive Knowledge Archive Network) is open-source data portal platform aimed at data publishers such as national and regional governments (including the U. S. government), companies and organizations wanting to make their data open and available. CKANs harvesting framework can be used to retrieve, normalize, and convert dataset metadata from multiple catalogs. It provides a catalog system, integration with third-party content management systems like Drupal and WordPress, data visualization and analytics, integrated data storage and full data API, and more. CKAN is maintained by the Open Knowledge Foundation which provides support and hosting.
- [DataHub](#)  
DataHub is a free data management platform from the Open Knowledge Foundation. It can be used to publish or register datasets as well as create and manage groups and communities. It is based on the CKAN data management system.
- [Dataverse Project](#)  
Dataverse is an open source web application to share, preserve, cite, explore, and analyze research data. It provides for the sharing and replication of data while providing academic credit and visibility for researchers, journals, data authors, publishers, data distributors, and affiliated institutions.
- [eXistdb](#)  
eXistdb is a NoSQL XML and non-documents database which uses the XML Query Language (XQuery) for coding and indexing. It can work alongside oXygen. Users of eXistdb include the Office of the Historian, United States Department of State and the University of Victoria Humanities Computing and Media Centre.
- [Fedora](#)  
Fedora (Flexible Extensible Digital Object Repository Architecture) is a modular, open source repository platform for the management and dissemination of digital content, including curating research data throughout the research life cycle from beginning through preservation in a RDF environment. Fedora is being used for digital collections, e-research, digital libraries, archives, digital preservation, institutional repositories, open access publishing, document management, digital asset management, and more.
- [Jupyter](#)  
Jupyter is an open-source web application for creating and sharing documents containing live code, equations, visualizations and narrative text. It can be used for data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more. Jupyter supports over 40 programming languages, including Python, R, Julia, and Scala.
- [KriKri](#)  
KriKri is a Ruby on Rails open source engine for metadata aggregation, enhancement, and quality control developed by the Digital Library of America (DPLA) released under the MIT License. It works with Heiðrún, DPLA's metadata ingestion system. Features include: harvesting metadata from OAI-PMH providers; creating RDF metadata models, with specific support for the DPLA Metadata Application Profile; enrichments for mapped metadata, including date parsing and normalization, stripping and splitting on punctuation; parsing metadata and mapping to RDF graphs using a Domain Specific Language; and more.
- [OpenRefine](#)  
OpenRefine (formerly Google Refine) is a tool for working with data. Use it to clean data, transform data from one format into another, extend data with web services, and link it to databases such as Wikidata.

- [Samvera](#)  
Samvera (previously, Hydra), is an open source digital asset management framework. The system uses Ruby gem building blocks allowing for customization. Samvera instances can be cloned and adapted to local needs. Bundled solutions requiring fewer local resources or cloud-based, hosted versions include Avalon, Hyrax, and Hyku.
- [Wikibase](#)  
Wikibase was developed for Wikidata as an open source collection of applications and libraries for creating and sharing structured data as linked data entities and their relationships. It consists of a set of extensions to the MediaWiki software for storing and managing data (Wikibase Repository) and for embedding data on other wikis (Wikibase Client). Wikibase provides an editing interface for creating, updating, merging, and deleting item and property entities.

## Discovery Interfaces

---

### [Blacklight](#)

Blacklight is an open source, discovery interface platform framework for searching an Apache Solr index. Blacklight MARC provides library catalog enhancements, Spotlight enables the creation of feature rich websites for digital collections, and Geoblacklight provides for the discovery and sharing of geospatial data. Search box, facet constraints, stable document urls, and more are customizable via Rails templating mechanisms. It accommodates heterogeneous data, allowing different information displays for different types of objects.

## Editors

---

- [Geany](#)  
Geany is an open source text editor using the GTK+ toolkit with basic features of an integrated development environment (IDE). It supports many filetypes including C, Java, Java Script, PHP, HTML, CSS, Python, Perl, Pascal, Ruby, XML, SQL, and more. Features include syntax highlighting, code folding, symbol name auto-completion, auto-closing of XML and HTML tags, code navigation, build system to compile and execute your code, symbol lists, and a plug-in interface. Geany runs on every platform which is supported by the GTK libraries including Linux, FreeBSD, NetBSD, OpenBSD, MacOS X, AIX v5.3, Solaris Express and Windows. Only the Windows port of Geany is missing some features.
- [LIME](#)  
Palmirani, Monica, Vitali, Fabio, and Cervone, Luca, et al. LIME is an open source, customizable web based editor for converting non-structured legal documents into XML. Currently, there are demo versions of LIME for three schema languages: AkomaNtoso; TEI; and LegalRuleML. LIME provides a linked outline view of the document and a contextual markup menu showing available elements. Click on the Demo tab at the top of the web site to choose a schema. LIME is under development at CIRSFD and the University of Bologna.
- [MarcEdit](#)  
MarcEdit is a free Marc editing tool. Use the tool to download a MARC record and transform it into an RDF/XML serialization of the record. The tool also can be used to perform MARC database maintenance. MarcEdit includes a tool for querying registered xslt crosswalks and downloading them for use with MarcEdit.
- [Notepad ++](#)  
Notepad ++ is a free source code editor that runs in the MS Windows environment.

- [oXygen](#)  
oXygen is a licensed cross platform XML editor that works with all XML-based technologies including XML databases, XProc pipelines, and web services. oXygen XML Author comes with a configurable and extensible visual editing mode based on W3C CSS stylesheets with ready-to-use DITA, DocBook, TEI, XHTML, XSLT, and XQuery support.
- [pymarc](#)  
Python Software Foundation. (2019). pymarc is a python library for working with bibliographic data encoded in MARC21. It provides an API for reading, creating, and modifying MARC records.
- [RDFa Play](#)  
RDFa Play is a real-time RDFa 1.1 editor, data visualizer and debugger. Paste your HTML+RDFa code into the editor to view a preview page, a data visualization, and the raw data of your code.

## Generators

---

### [Dublin Core Generator](#)

This site provides three tools developed by Nick Steffel to generate Dublin Core code. The Simple Generator generates simple Dublin Core metadata using only the 15 main elements. Advanced Dublin Core metadata code using the more detailed qualified elements and encoding schemes can be generated using the Advanced Generator, and there is a generator for the xZINECOREx variation of Dublin Core.

## Glossaries

---

- [Glossary of Metadata Standards](#)  
This glossary lists the most common metadata standards used in the cultural heritage community. Several of them are listed on our Vocabularies page, which you can access by clicking on Vocabularies, etc. in the menu on the left. A color version of the Seeing Standards poster is also shown on that page. A poster version of the glossary is also available.
- [Glossary of Terms Relating to Thesauri and Other Forms of Structured Vocabulary](#)  
Will, Leonard D. and Will, Sheena. (2013). This is an alphabetical list of terms associated with thesauri and structured vocabularies.
- [Linked Data Glossary](#)  
This is the W3C's glossary of Linked Data terms.

## Ontology/Vocabulary Building Platforms and Tools

---

- [Fluent Editor](#)  
Fluent Editor is a tool for editing and manipulating complex ontologies that use Controlled Natural Language. A main feature is the usage of Controlled English as a knowledge modeling language. it prohibits one from entering any sentence that is grammatically or morphologically incorrect and actively helps the user during sentence writing. It is free for individual or academic use. Access to updates and information is given with registration.
- [Neologism](#)  
Neologism is an open source vocabulary publishing platform for creating and publishing vocabularies compatible with Linked Data principles. It supports the RDFS standard enabling you to create RDF classes and properties. It also supports a part of OWL. Neologism is written in PHP and built on the Drupal platform.

- [NeOn](#)  
NeOn is an open source multi-platform for the support of the ontology engineering life-cycle. The toolkit is based on the Eclipse platform and provides an extensive set of plug-ins covering a variety of ontology engineering activities, including Annotation and Documentation, Development, Human-Ontology Interaction, Knowledge Acquisition, Management, Modularization and Customization, Neon Plugins, Old Main Page, Ontology Dynamics, Ontology Evaluation, Ontology Matching, Reasoning and Inference, and Reuse. NeOn's aim is to advance the state of the art in using ontologies for large-scale semantic applications in distributed organisations by improving the ability to handle multiple networked ontologies that exist in a particular context, are created collaboratively, and might be highly dynamic and constantly evolving.
- [OOPS! \(OntOlogy Pitfall Scanner!\)](#)  
OOPS! is an application used to detect common pitfalls when developing ontologies. Enter the URI or the RDF code of the ontology. Once the ontology is analyzed, a results list of pitfalls appear that can be expanded to display information regarding the pitfalls.
- [Protégé](#)  
Protégé is a free, opensource platform with a suite of tools to construct domain models and knowledge - based applications with ontologies. Protégé Desktop is a feature rich ontology editing environment with full support for the OWL 2 Web Ontology Language, and direct in-memory connections to description logic reasoners like HermiT and Pellet. Protégé Desktop supports creation and editing of one or more ontologies in a single workspace via a completely customizable user interface. Visualization tools allow for interactive navigation of ontology relationships. It is W3C standards compliant and offers ontology refactoring support, direct interface to reasoners like HermiT and Pellet and is cross compatible with WebProtégé. Protégé provides an environment to create, upload, modify, and share ontologies for collaborative viewing and editing. Protégé was developed by the Stanford Center for Biomedical Informatics Research at the Stanford University School of Medicine. Download the desktop version or use the Web version from this site.

#### VOWL: Visual Notation for OWL Ontologies

This page provides access to three tools for visualizing ontologies: WebVOWL; QueryVOWL; and the Protégé plug-in, ProtégéVOWL. A link to the VOWL (Visual Notation for OWL Ontologies) specification and a Language Reference for QueryVOWL (Visual Query Language) for Linked Data is also provided.

## Query Tools, Search Engines & Browser Add-ons

---

- [Linked Data Fragments](#)  
Use this tool to execute queries against live Linked Data on the Web in your browser. The tool supports federated querying.
- [OpenLink Data Explorer Extension](#)  
OpenLink Software. This web browser extension provides options for viewing Data Sources associated with Web Pages to explore the raw data and entity relationships that underlay the Web resources it processes. The extension enables Hypertext and Hyperdata traversal of Web data. The browser add-on is easy to install. It was first developed for use on most browsers, but with some browser updates, the add-on doesn't work. Try using it with Chrome. The browser provides filters for faceted searching and visualization options.

- [OpenLink Structured Data Sniffer \(OSDS\)](#)  
OpenLink Software. OpenLink Structured Data Sniffer is a browser extension for Google Chrome, Microsoft Edge, Mozilla Firefox, Opera, and Vivaldi that reveals structured metadata embedded in HTML pages in notations including POSH (Plain Old Semantic HTML), Microdata, JSON-LD, RDF-Turtle, and RDFa. Buttons assist in navigating the Web, and it provides the ability to save extracted metadata or new annotations to the cloud or local storage.

## Servers

---

- [Metaproxy](#)  
Index Data. Metaproxy is a proxy Z39.50/SRW/SRU front end server designed for integrating multiple back end databases into a single searchable resource. It also works in conjunction with Index Data's library of gateways to access non-standard database servers. Index Data works with libraries, consortia, publishers, aggregators, technology vendors, and developers.
- [Ontobee](#)  
He Group. University of Michigan. Ontobee is a Linked Data server designed to facilitate ontology sharing, visualization, query, integration, and analysis. It dereferences term URIs to HTML web pages for user-friendly browsing and navigation and to RDF source code for Semantic Web applications.

## Triple Store Tools

---

- [Blazegraph](#)  
Blazegraph is a scalable, high-performance graph database with support for Blueprints and RDF/SPARQL APIs. It supports up to 50 billion edges on a single machine. Blazegraph works in a Python environment. Wikimedia uses it to power their wikidata query service.
- [Gruff](#)  
Gruff is a free, downloadable graphical triple-store browser with a variety of tools for laying out cyclical graphs, displaying tables of properties, managing queries, and building queries as visual diagrams. Use gruff to display visual graphs of subsets of a store's resources and their links and build a visual graph that displays a variety of the relationships in a triple-store. Gruff can also display tables of all properties of selected resources or generate tables with SPARQL queries, and resources in the tables can be added to the visual graph.

## Validators

---

- [Freeformatter JSON Validator](#)  
This tool validates a JSON string against RFC 4627 (the application/json media type for JavaScript Object Notation) and against the JavaScript language specification. Configure the validator to be lenient or strict.
- [Link Checker](#)  
W3C. (2019). Use this validator to check issues with links, anchors and referenced objects in Web pages, CSS style sheets, or whole Web sites. Best results are achieved when the documents checked use Valid (X)HTML Markup and CSS.
- [RDF Validation Service](#)  
Use this tool to parse RDF/XML documents. A 3-tuple (triple) representation of the corresponding data model as well as an optional graphical visualization of the data model will be displayed.



- [Structured Data Linter](#)  
The Structured Data Linter was initiated by Stéphane Corlosquet and Gregg Kellogg. It is a tool to verify structured data present in HTML pages. The Linter provides snippet visualizations for schema.org and performs limited vocabulary validations for schema.org, Dublin Core Metadata Terms, Friend of a Friend (FOAF), GoodRelations, Facebook's Open Graph Protocol, Semantically-Interlinked Online Communities (SIOC), Facebook's Open Graph Protocol, Simple Knowledge Organization System (SKOS), and Data-Vocabulary.org.
- [Toolz Online XML Validator](#)  
Insert a fragment of an XML document into this tool to validate it.
- [Yandex](#)  
Yandex is a structured data Microformat validator for checking semantic markup. Check all the most common microformats: microdata, schema.org, microformats, OpenGraph and RDF by cutting and pasting the source code into the validator.

## Visualization Tools

---

- [D3 Data-Driven Documents](#)  
is a JavaScript library for manipulating documents based on data using HTML, SVG and CSS. Using D3, data can be displayed in a vast array of visualization formats including, but not limited to Box Plots, Bubble Charts, Bullet Charts, Calendar Views, Chord Diagrams, Dendograms, Force-Directed Graphs, Chord Diagrams, Circle Packings, Population Pyramids, Steamgraphs, Sunbursts, Node-link Trees, Treemaps, Voronoi Diagrams, Collision Detections, Hierarchical Edge Bundlings, Word Cloud, and more.
- [Visual Data Web](#)  
The Visual Data Web provides links to visualization tools compatible with RDF and Linked Data on the Semantic Web, especially for average Web users with little to no knowledge about the underlying technologies. The site provides information regarding developments, related publications, and current activities to generate new ideas, methods, and tools to make the Data Web more accessible and visible.

## XPath Tools

---

- [eagle-i](#)  
The eagle-i Software and ontology consists of six web applications: eagle-i Central Search and iPS Cell Search — for resource discovery and exploration; Institutional search — for a single repository search UI; Ontology Browser — for viewing the eagle-i ontology without any additional applications; SWEET (Semantic Web Entry & Editing Tool) — for manually entering and managing data in an eagle-i repository; RDF repository — for storing resource and provenance metadata as RDF triples; and SPARQLer — a SPARQL query entry point and workbench to query an eagle-i repository. These applications are served by the ETL (extract, transform, and load) toolkit — for batch entry of information to an eagle-i repository in an ontology-compliant manner and the Data management toolkit — for bulk data maintenance and migration. The open source software development platform offers integrated tools for JIRA bug tracking, Confluence Wiki, Bamboo continuous builds, Nexus download repository, project mailing lists, repository monitoring, and more.
- [Freeformatter XPath Tester/Evaluator](#)  
Use this tool to test XPath expressions/queries against an XML file. It supports most of the XPath functions (string(), number(), name(), string-length() etc.) and is not limit to working against nodes.
- [Toolz XPath Tester/Evaluator](#)  
Use this tool to run an XPATH statement against an XML fragment

- [W3C XPath evaluation online](#)  
Use this W3C tool to check a XPath expression against XML.

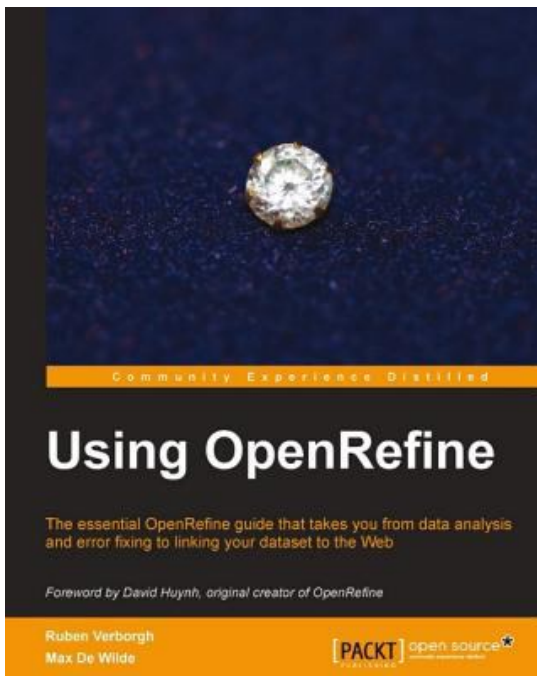
## Miscellaneous

---

- [Keyword Planner](#)  
The Google AdWords Keyword Planner is not a semantic web tool. While geared towards advertising, it can be a useful tool to discover similar keywords for a topic. It is a free tool, but you will have to create an account.
- [prefix.cc](#)  
Enter a namespace prefix in this tool to find the full namespace for the prefix. The service also provides a reverse lookup option which finds a prefix for a given namespace URI.

## Instructional Resources for Semantic Tools

---



[Using Openrefine by Ruben Verborgh; Max De Wilde](#)  
ISBN: 9781783289097

Date: 2013-09-10

This book describes how to use the browser plug-in, OpenRefine, to clean data, transform data from format to format, link data to databases such as Freebase, and extend data with web services.

- [WebLearn](#)  
This blog provides examples of using OpenRefine to clean MARC data. Stephen shares his experience working with MARC data while developing the Sir Louie Project, a project to improve the searching of library catalogues and the displaying availability information with a reading list on behalf of the British Library.
- [MARCEdit You Tube Videos](#)  
This page lists over 90 videos produced by Terry Reese providing instructions for using MARCEdit. Topics include "MarcEdit 101: I have a MARC record, now what?," "Installing MarcEdit natively on a Mac operating system," "Extract and Edit Subsets of Records in MarcEdit," "MarcEdit Task Automation Tool," and "MarcEdit RDA Helper."

## SPARQL

---

## SPARQL

---

SPARQL serves as the search engine for RDF. It is a set of specifications recommended by W3C Recommendation that provide languages and protocols to query and manipulate RDF graph content on the Web or in an RDF triple store.

### SPARQL Documentation

- [SPARQL 1.1 Entailment Regimes](#)  
Glimm, Birte, and Ogbuji, Chimezie, editors. (2013, March 21). This document defines entailment regimes and specifies how they can be used to redefine the evaluation of basic graph patterns from a SPARQL query making use of SPARQL's extension point for basic graph pattern matching. Entailment regimes specify conditions that limit the number of entailments that contribute solutions for a basic graph pattern.
- [SPARQL 1.1 Federated Query](#)  
Seaborne, Andy, Polleres, Axel, Feigenbaum, Lee, and Williams, Gregory Todd. (2013, March 21). The SPARQL Federated Query extension is a specification which defines the syntax and semantics for using the SERVICE keyword to execute queries that merge data distributed over different SPARQL endpoints. It provides for the ability to direct a portion of a query to a particular SPARQL endpoint. Results are returned to the federated query processor and are combined with results from the rest of the query.
- [SPARQL 1.1 Graph Store HTTP Protocol](#)  
Ogbuji, Chimezie, editor. (2013, March 21). This document describes the use of HTTP for managing a collection of RDF graphs as an alternative to the SPARQL 1.1 Update protocol interface. For some clients or servers, HTTP may be easier to implement or work with, and this specification serves as a non-normative suggestion for HTTP operations on RDF graphs which are managed outside of a SPARQL 1.1 graph store.
- [SPARQL 1.1 Overview](#)  
W3C SPARQL Working Group. (2013, March 21). This document provides an introduction to a set of W3C specifications for querying and manipulating RDF graph content on the Web or in an RDF store. It gives a brief description of the eleven specifications that comprise SPARQL.
- [SPARQL 1.1 Protocol](#)  
Feigenbaum, Lee, Williams, Gregory Todd, Clark, Kendall Grant, Torres, Elias. (2013, March 21). The SPARQL 1.1 Protocol describes a means for conveying SPARQL queries and updates to a SPARQL processing service and returning the results via HTTP to the entity that requested them. It has been designed for compatibility with the SPARQL 1.1 Query Language [SPARQL] and with the SPARQL 1.1 Update Language for RDF. The intended use of this document is primarily intended for software developers implementing SPARQL query and update services and clients.
- [SPARQL 1.1 Query Results CSV and TSV Formats](#)  
Seaborne, Andy. (2013, March 21). This document describes the use of Comma Separated Values (CSV) and tab separated values (TSV) for expressing SPARQL query results from SELECT queries. CSV and TSV are formats for the transmission of tabular data, particularly spreadsheets.
- [SPARQL 1.1 Query Results JSON Format](#)  
Clark, Kendall Grant, Feigenbaum, Lee, Torres, Elias. (2013, March 21). SPARQL is a set of standards which defines several Query Result Forms used to query and update RDF data, along with ways to access such data over the web. This document defines the representation of SELECT and ASK query results using JSON.

- [SPARQL Query Results XML Format \(Second Edition\)](#)  
Beckett, Dave, and Broekstra, Jeen. (2013, March 21). SPARQL is a set of standards which defines several Query Result Forms used to query and update RDF data, along with ways to access such data over the web. This document defines the SPARQL Results Document that encodes variable binding query results from SELECT queries and boolean query results from ASK queries in XML.
- [SPARQL 1.1 Query Language](#)  
Harris, Steve, Seaborne, Andy. (2013, March 21). This document defines the syntax and semantics of the SPARQL query language for RDF, a directed, labeled graph data format for representing information in the Web. SPARQL is used to express queries across data sources, whether the data is stored natively as RDF or viewed as RDF via middleware. SPARQL supports querying required and optional graph patterns along with their conjunctions and disjunctions, aggregation, subqueries, negation, creating values by expressions, extensible value testing, and constraining queries by source RDF graph. Results of SPARQL queries can be result sets or RDF graphs.
- [SPARQL 1.1 Service Description](#)  
Williams, Gregory Todd. (2013, March 21). A SPARQL service description lists the features of a SPARQL service made available via the SPARQL Protocol for RDF. This document describes how to discover a service description from a specific SPARQL service and an RDF schema for encoding such descriptions in RDF.
- [SPARQL 1.1 Update](#)  
Gearon, Paula, Passant, Alexandre, and Polleres, Axel. (2013, March 21). SPARQL 1.1 Update is a language used to update RDF graphs using a syntax derived from the SPARQL Query Language for RDF. Operations are provided to update, create, and remove RDF graphs in a Graph Store.

## GeoSpatial SPARQL

In addition to the W3c SPARQL documents, there is documentation for a Geospatial SPARQL query language.

### [OGC GeoSPARQL - A Geographic Query Language for RDF Data](#)

Matthew Perry, Matthew, and Herring, John, editors. (2012, September 10). Open Geospatial Consortium (OGC). This OGC standard defines a vocabulary for representing geospatial data in RDF. It also defines an extension to the SPARQL query language for processing geospatial data. The GeoSPARQL query language is designed to accommodate systems based on qualitative spatial reasoning and systems based on quantitative spatial computations.

## SPARQL Endpoints

---

This box provides links to some SPARQL endpoints that are useful for researchers, and are good examples of datasets to practice using SPARQL queries. The Europeana dataset is used in the SPARQL for humanists tutorial on the left.

### [Europeana SPARQL API](#)

Use this API to explore connections between Europeana data and outside data sources, like VIAF, Iconclass, Getty Vocabularies (AAT), Geonames, Wikidata, and DBPedia.

## SPARQL Tools

---

This box contains SPARQL tools.

- [Apache Jena Fuseki2](#)  
Apache Jena Fuseki is a SPARQL server. It has the capability to run as an operating system service, as a Java web application (WAR file), and as a standalone server. It provides SPARQL 1.1 protocols for query, update and the SPARQL Graph Store. Fuseki can be configured with TDB to provide a transactional persistent storage layer, and incorporates Jena text query and Jena spatial query.
- [Pubby](#)  
Bizer, Christian, and Cyganiak, Rhichard. Freie Universität Berlin. Pubby adds Linked Data interfaces to SPARQL endpoints. It can turn a SPARQL endpoint into a Linked Data server, and is implemented as a Java web application. Features include providing dereferenceable URIs by rewriting URIs found in the SPARQL-exposed dataset into the Pubby server's namespace, providing an HTML interface showing the data available about each resource, handling 303 redirects and content negotiation, and provides for the addition of metadata. It is compatible with Tomcat and Jetty servlet containers.

## SPARQL Instructional Resources

---

- [SPARQL Sample Queries](#)  
Coombs, Karen. This page on the github blog, Library Web Chic, provides useful examples of SPARQL queries. This is an excellent place to browse through when learning how to query with SPARQL. Examples include simple queries for finding subjects, predicates, and objects and build into more complex federated and filtered queries across datasets. This serves as a companion to Karen Coombs' Querying Linked Data webinar.
- [SPARQL for humanists](#)  
Lincoln, Matthew. (2014, July 10). From The Programming Historian. This blog entry describes using SPARQL using the Europeana Data Model (EDM). It provides a good introduction to SPARQL.  
  
[more...](#)
- [Using SPARQL to access Linked Open Data](#)  
Lincoln, Matthew. (2015, November 24). From The Programming Historian. This blog entry provides a lesson explaining why cultural institutions are moving to graph databases. The entry also gives a detailed lesson in using SPARQL to access data in cultural institution databases.

## Vocabularies, Ontologies & Frameworks

---

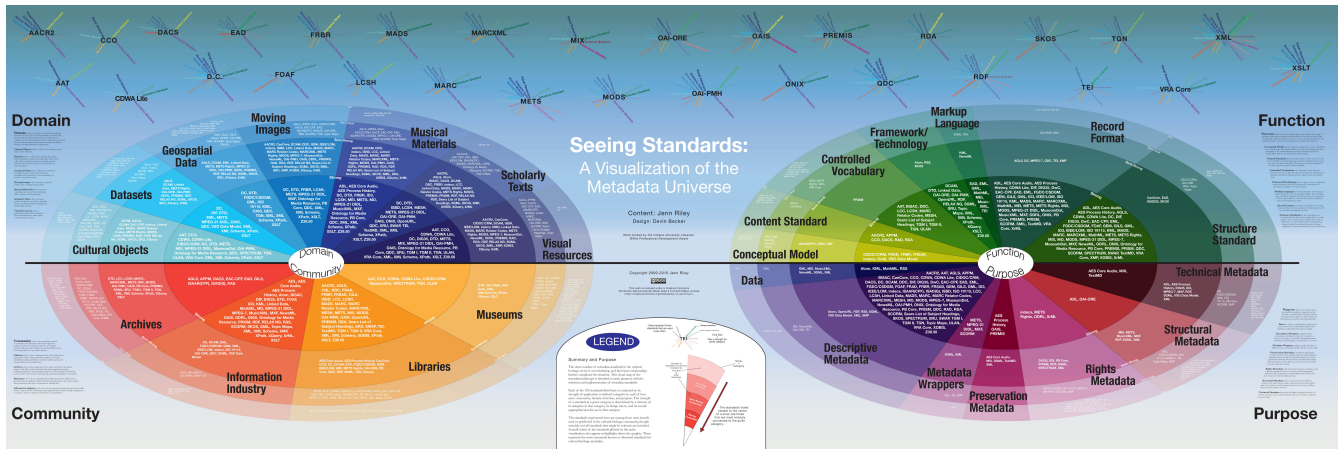
### Vocabularies, Ontologies & Frameworks

---

Controlled vocabularies, ontologies, schemas, thesauri, and syntaxes are building blocks used by Resource Description Framework (RDF) to structure data semantically, identify resources, and to show the relationships between resources in Linked Data. Libraries and cultural institutions belong to one of the many knowledge organization domains making use of controlled authorities. These pages focus especially on the vocabularies and computer languages that are used in the library and cultural heritage institutions data landscape.

## Seeing Standards: Visualization of the Metadata Universe

---



## About Seeing Standards

Becker, Devin and Jenn L. Riley. (2010). *Seeing Standards: A Visualization of the Metadata Universe*. Click on the chart to access a PDF version and a Glossary of Metadata Standards.

## About Vocabularies

### About Taxonomies & Controlled Vocabularies

American Society for Indexing, Taxonomies & Controlled Vocabularies Special Interest Group. This page describes the differences between controlled vocabularies, taxonomies, thesauri, and ontologies.

## Ontologies and Frameworks

## Ontologies & Frameworks

### International Image Interoperability Framework (IIIF)

IIIF is a framework for image delivery developed by a community of leading research libraries and image repositories. The goals are to provide access to an unprecedented level of uniform and rich access to image-based resources hosted around the world, define a set of common application programming interfaces supporting interoperability between image repositories, develop, cultivate and document shared technologies, such as image servers and web clients, for providing viewing, comparing, manipulating, and annotating images.

The two core APIs for the Framework are:

#### IIIF Image API 3.0

IIIF (International Image Interoperability Framework) Consortium. (2021). Appleby, Michael, Crane, Tom, Sanderson, Robert, Stroop, Jon, and Warner, Simeon. This document describes an image delivery API specification for (Application Programming Interface) a web service that returns an image in response to a standard HTTP (Hypertext Transfer Protocol) or HTTPS (Hypertext Transfer Protocol (Secure)) request. The URI can specify the region, size, rotation, quality characteristics and format of the requested image as well as be enabled to request basic technical information about the image to support client applications.

#### IIIF Presentation API 3.0.

IIIF Consortium. (2021). Appleby, Michael, Crane, Tom, Sanderson, Robert, Stroop, Jon, and Warner, Simeon. The IIIF Presentation API.(Application Programming Interface)provides information necessary to human users to allow a rich, online viewing environment for compound digital objects. It enables the display of digitized images, video, audio, and other content types associated with a particular physical or born-digital object, allows navigation between multiple views or time extents of the object, either sequentially or hierarchically, displays descriptive information about the object, view or navigation structure, and provides a shared environment in which publishers and users can annotate the object and its content with additional information.

### Presentation Cookbook of IIIF Recipes

The Cookbook provides resource types and properties of the Presentation specification and for rendering by viewers and other software clients. Examples are provided to encourage publishers to adopt common patterns in modeling classes of complex objects, enable client software developers to support these patterns, for consistency of user experience, and demonstrate the applicability of IIIF to a broad range of use cases.

Additional APIs for the Framework are:

### IIIF Authentication API 1.0

IIIF Consortium. (2021). Appleby, Michael, Crane, Tom, Sanderson, Robert, Stroop, Jon, and Warner, Simeon. The Authentication specification describes a set of workflows for guiding the user through an *existing* access control system. It provides a link to a user interface for logging in, and services that provide credentials, modeled after elements of the OAuth2 workflow acting as a bridge to the access control system in use on the server, without the client requiring knowledge of that system.

### IIIF Content Search API 1.0

IIIF Consortium. (2021). Appleby, Michael, Crane, Tom, Sanderson, Robert, Stroop, Jon, and Warner, Simeon. The Content Search specification lays out the interoperability mechanism for performing searches among varied content types from different sources. The scope of the specification is searching annotation content within a single IIIF.(International Image Interoperability Framework) resource, such as a Manifest, Range or Collection.

## **Linked Art**

**Linked Art is a data model which provides an application profile used to describe cultural heritage resources, with a focus on artworks and museum-oriented activities. Based on real world data and use cases, it** defines common patterns and terms used in its conceptual model, ontologies, and vocabulary. Linked Art follows existing standards and best practices including CIDOC-CRM, Getty Vocabularies, and JSON-LD 1.1 as the core serialization format.

## **OWL 2**

Ontologies are formalized vocabularies of terms, often covering a specific domain. They specify the definitions of terms by describing their relationships with other terms in the ontology. OWL 2 is the Web Ontology Language designed to facilitate ontology development and sharing via the Web. It provides classes, properties, individuals, and data values that are stored as Semantic Web documents. As an RDF vocabulary, OWL can be used in combination with RDF schema.

### **VOWL: Visual Notation for OWL Ontologies**

Negru,Stefan, Lohmann, Seffan, and Haag, Florian. (2014, April 7). Specification of Version 2.0. VOWL defines a visual language for user-oriented representation of ontologies. The language provides graphical depictions for elements of OWL that are combined to a force-directed graph layout visualizing the ontology. It focuses on the

visualization of the classes, properties and datatypes, sometimes called TBox, while it also includes recommendations on how to depict individuals and data values, the ABox. Familiarity with OWL and other Semantic Web technologies is required to understand this specification.

- [OWL 2 Web Ontology Language Document Overview \(Second Edition\)](#)  
This is the W3C introduction to OWL 2 and the various other OWL 2 documents. The document describes the syntaxes for OWL 2, the different kinds of semantics, the available sub-languages, and the relationship between OWL 1 and OWL 2. Read this document before reading other OWL 2 documents.
- [OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax \(Second Edition\)](#)  
This document defines the OWL 2 language. The core part, the structural specification, describes the conceptual structure of OWL 2 ontologies and provides a normative abstract representation for all OWL 2 syntaxes. The document also defines the functional-style syntax, which follows the structural specification and allows OWL 2 ontologies to be written in a compact form. This syntax is used in the definitions of the semantics of OWL 2 ontologies, the mappings from and into the RDF/XML exchange syntax, and the different OWL 2 profiles.
- [OWL 2 Web Ontology Language Mapping to RDF Graphs \(Second Edition\)](#)  
This document defines two mappings between the structural specification of OWL 2 and RDF graphs. The mappings can be used to transform an OWL 2 ontology into an RDF graph and an RDF graph into an OWL 2 ontology.
- [Time Ontology in Owl](#)  
Cox, Simon, Little, Chris, Hobbs, Jerry R., and Pan, Feng. (2017, October 19). W3C. Time Ontology in Owl (OWL-Time) can be used to describe temporal relationships. It focuses particularly on temporal ordering relationships. Elements of a date and time are put into separately addressable resources. OWL-Time supports temporal coordinates (scaled position on a continuous temporal axis) and ordinal times (named positions or periods) and does not necessarily expect the use of the Gregorian calendar.

## **PRESSoo**

### PRESSoo

Le Boeuf, Patrick (2016, January). PRESSoo is an ontology designed to represent bibliographic information relating to serials and continuing resources. PRESSoo is an extension of the Functional Requirements for Bibliographic Records – Object Oriented model (FRBRoo). PRESSoo has been developed by representatives of the ISSN International Centre, the ISSN Review Group, and the Bibliothèque nationale de France (BnF).

## **RDF 1.1**

The Resource Description Framework (RDF) is a framework for representing information in the Web of Data. It comprises a suite of standards and specifications whose documentation is listed below.

- [Cool URIs for the Semantic Web](#)  
Leo Sauermann, Leo and Cyganiak, Richard. (2008, Dec.3). W3C.  
Uniform Resource Identifiers (URIs) are at the core of RDF providing the link between RDF and the Web. This document presents guidelines for their effective use. It discusses two strategies, called 303 URIs and hash URIs. It gives pointers to several Web sites that use these solutions, and briefly discusses why several other proposals have problems.



- [RDF 1.1 Concepts and Abstract Syntax](#)  
This W3C document defines an abstract syntax (a data model) for linking RDF-based languages and specifications. The syntax has a data structure for expressing descriptions of resources as RDF graphs made of sets of subject-predicate-object triples, where the elements may be IRIs, blank nodes, or datatyped literals. The document introduces key concepts and terminology, and discusses datatyping and the handling of fragment identifiers in IRIs within RDF graphs.
- [RDF 1.1 Primer](#)  
The Primer introduces basic RDF concepts and shows concrete examples of the use of RDF. It is designed to provide the basic knowledge required to effectively use RDF.
- [RDF Schema 1.1](#)  
The RDF Schema provides a data-modeling vocabulary for RDF data and is an extension of the basic RDF vocabulary. The IRIs for the namespaces for the RDF Schema and the RDF Syntax are defined in this document. The RDF Schema provides mechanisms for describing groups of related resources and the relationships between these resources which can be used to describe other RDF resources in application-specific RDF vocabularies.
- [RDF 1.1 Semantics](#)  
This is one of the documents that comprise the full specification of RDF 1.1. It describes semantics for the Resource Description Framework 1.1, the RDF Schema, and RDFS vocabularies.

## **RDF 1.1 Serializations**

There are a number of RDF serialization formats for implementing RDF. The first format was XML/RDF. Subsequent serialization formats have been developed and may be more suited to particular environments.

- [JSON-LD 1.0](#)  
Sporny, Manu, Longley, Dave, Kellogg, Gregg, Lanthaler, Markus, and Lindström, Niklas. (2014, Jan.16).A JSON-based Serialization for Linked Data. Recommendation. W3C.  
This specification defines JSON-LD, a JSON-based format to serialize Linked Data. JSON-LD with RDF tools can be used as a RDF syntax.
- [RDF 1.1 Turtle](#)  
Terse RDF Triple Language.  
David Beckett, Berners-Lee, Tim, Prud'hommeaux, Eric, and Carothers, Gavin. (2014, Feb.25). Recommendation. W3C.  
This document defines Turtle, the Terse RDF Triple Language, a concrete syntax for RDF that allows an RDF graph to be written in a compact, natural text form with abbreviations for common usage patterns and datatypes. Turtle provides levels of compatibility with the N-Triples format and SPARQL.
- [RDF 1.1 XML Syntax](#)  
This W3C document defines the XML syntax for RDF graphs.  
W3C. (2014, Feb.25). Recommendation. Gandon, Fabien and Schreiber, Guus. eds.
- [RDFa Core 1.1](#)  
Adida, Ben, Birbeck, Mark, McCarron, Shane, and Herman, Ivan. (2015, Mar. 17).  
Syntax and processing rules for embedding RDF through attributes. 3rd. ed. Recommendation. W3C.  
RDFa Core is a specification for attributes to express structured data in any markup language. The rules for interpreting the data are generic, so that there is no need for different rules for different formats. The embedded data already available in the markup language (e.g., HTML) can often be reused by the RDFa markup

- [RDFa 1.1 Primer](#)  
Herman, Ivan, Adida, Ben, Sporny, Manu, and Birbeck, Mark. (2015, Mar. 17). Rich Structured Data Markup for Web Documents. W3C. RDFa (Resource Description Framework in Attributes) is a technique to add structured data to HTML pages directly. This Primer shows how to express data using RDFa in HTML, and in particular how to mark up existing human-readable Web page content to express machine-readable data.

## **SKOS (Simple Knowledge Organization System)**

SKOS is a W3C data model defined as an OWL Full ontology for use with knowledge organization systems including thesauri, classification schemes, subject heading systems, and taxonomies. Many Semantic Web vocabularies incorporate the SKOS model. The Library of Congress Subject Headings and the Getty Vocabularies are examples of vocabularies published as SKOS vocabularies.

- [SKOS Simple Knowledge Organization System eXtension for Labels \(SKOS-XL\) Namespace Document - HTML Variant](#)  
SKOS-XL defines an extension for SKOS which provides additional support for describing and linking lexical entities. This document provides a brief description of the SKOS-XL vocabulary.
- [SKOS Simple Knowledge Organization System Namespace Document - HTML Variant](#)  
This W3C document provides an HTML non-normative table of the SKOS vocabulary.
- [SKOS Simple Knowledge Organization System Primer](#)  
SKOS provides a model for expressing the basic structure and content of concept schemes such as thesauri, classification schemes, subject heading lists, taxonomies, folksonomies, and other similar types of controlled vocabulary. This document serves as a user guide for those who would like to represent their concept scheme using SKOS.
- [SKOS Simple Knowledge Organization System Reference](#)  
This document defines the SKOS namespace and vocabulary. SKOS is a data-sharing standard which aims to provide a bridge between different communities of practice within the library and information sciences involved in the design and application of knowledge organization systems widely recognized and applied in both modern and traditional information systems.

## **Ontology Development**

---

[Ontology Development 101: A Guide to Creating Your First Ontology](#)

Noy, Natalya F. and McGuinness, Deborah L. This guide discusses the reasons for developing an ontology and the methodology for creating an ontology based on declarative knowledge representation systems.

## **Registries, Portals, and Authorities**

---

## **Open Linked Vocabularies (LOV)**

---



- [Linked Open Vocabularies \(LOV\)](#)  
Use this site to find a list of vetted linked open vocabularies (RDFS or OWL ontologies) used in the Linked Open Data Cloud, which conform to quality requirements including URI stability and availability on the Web, use of standard formats and publication best practices, quality metadata and documentation, an identifiable and trusted publication body, and proper versioning policy. Vocabularies are individually described by metadata and classified by the following vocabulary spaces: General and Meta; Library; City; Market; Space-Time; Media; Science; and Web. They are interlinked using the dedicated vocabulary VOF. Search the LOV dataset at the vocabulary or element level. LOV Stats provide metric information regarding the vocabularies such as the number of vocabulary element occurrences in the LOD, the number of vocabularies in LOV that refer to a particular element, and more.
- [Open Metadata Registry](#)  
The Registry provides a means to identify, declare, and publish through registration metadata schemas (element/property sets), schemes (controlled vocabularies) and Application Profiles (APs). It supports the machine mapping of relationships among terms and concepts in those schemes (semantic mappings) and schemas (crosswalks). The Registry supports metadata discovery, reuse, standardization, and interoperability locally and globally.
- [RDA Registry](#)  
The RDA Registry defines vocabularies that represent the Resource Description Access (RDA) element set, relationship designators, and controlled terminologies as RDA element sets and RDA value vocabularies in Resource Description Framework (RDF). The published vocabularies are currently available in several sets which reflect the underlying FRBR conceptual model.
- [TaxoBank Terminology Registry](#)  
TaxoBank contains information about controlled vocabularies of all types and complexities. The information collected about each vocabulary follows a study conducted by the Joint Information Systems Committee (JISC) of the Higher and Further Education Funding Councils. The site offers additional resources including information on Thesauri and Vocabulary Control - Principles and Practice, and a Glossary of Terms Relating to Thesauri and Other Forms of Structured Vocabulary.
- [Virtual International Authority File \(VIAF\)](#)  
VIAF is a utility that matches and links authority files of national libraries. Data are derived from the personal name authority and related bibliographic data of the participating libraries. VIAF is implemented and hosted by OCLC.

## The Getty Vocabularies

---

- [Art & Architecture Thesaurus® Online](#)  
The Getty Research Institute. The scope of this vocabulary includes terminology needed to catalog and retrieve information about the visual arts and architecture
- [Cultural Objects Name Authority® Online and Iconography Authority \(IA\)](#)  
The Getty Research Institute. The Cultural Objects Name Authority ® (CONA) compiles titles, attributions, depicted subjects, and other metadata about works of art, architecture, and other cultural heritage, both extant and historical, physical and conceptual and can be used to record works depicted in visual surrogates or other works. Metadata may be gathered and linked from photo archive collections, visual resource collections, special collections, archives, libraries, museums, scholarly research, and other sources. The Getty Iconography Authority (IA) includes proper names and other information for named events, themes and narratives from religion/mythology, legendary and fictional characters, themes from literature, works of literature and performing arts, and legendary and fictional places.

- [The Getty Vocabularies as Linked Open Data](#)  
The Getty Art & Architecture Thesaurus (AAT) ®, Thesaurus of Geographic Names (TGN) ®, and the Union List of Artist Names (ULAN) ® are available as Linked Open Data. This link provides access the vocabularies and information regarding how to use them. Examples of URIs for each vocabulary are provided.
- [Getty Vocabularies: Linked Open Data Semantic Representation](#)  
Vladimir Alexiev, Joan Cobb, Gregg Garcia, Patricia Harpring. (2017, June 13). This document explains the representation of the Getty Vocabularies in semantic format, using RDF and appropriate ontologies. It covers the Art and Architecture Thesaurus (AAT)®, the Thesaurus of Geographic Names (TGN)® and the Union List of Artist Names (ULAN)®.
- [Getty Vocabularies OpenRefine Reconciliation](#)  
The Getty Research Institute. This page offers information and a tutorial on how to reconcile data sets to the Getty Vocabularies using the browser add-on OpenRefine. Use data reconciliation to compare local data to values in the Getty Vocabularies in order to map to them.
- [Thesaurus of Geographic Names® Online](#)  
The Getty Research Institute. The scope of this vocabulary spans a wide spectrum of geographic vocabulary in cataloging and scholarship of art and architectural history and archaeology.
- [Traing Materials](#)  
The Getty Research Institutes. This page provides training materials for the Art & Architecture Thesaurus (AAT)®, the Getty Thesaurus of Geographic Names (TGN)®, the Union List of Artist Names (ULAN)®, the Cultural Objects Name Authority (CONA)®, the Getty Iconography Authority (IA)™, Categories for the Description of Works of Art (CDWA), Cataloging Cultural Objects (CCO), and standards in general. It also provides access conference presentations.
- [Union List of Artist Names® Online](#)  
The Getty Research Institute. The ULAN is a structured vocabulary containing names and other information about artists, patrons, firms, museums, and others related to the production and collection of art and architecture. Names in ULAN may include given names, pseudonyms, variant spellings, names in multiple languages, and names that have changed over time (e.g., married names).

## Schemas

## Schemas

---

A schema uses a formal language to describe a database system and refers to how the organization of data in a database is constructed. Several schemas addressing varied domain areas are listed in this box. Scroll down to the Dublin Core box to access information regarding the Dublin Core schema and tools.

- [BIBFRAME \(Bibliographic Framework\) Initiative](#)  
This is the homepage for BIBFRAME the Library of Congress' Bibliographic Framework Initiative. BIBFRAME is a replacement for MARC and serves as a general model for expressing and connecting bibliographic data to the Web of Data. Access links to general information, the vocabulary, BIBFRAME implementation register, tools, draft specifications for Profiles, Authorities, and Relationships, a BIBFRAME testbed, webcasts and presentations, and more.

- [BIBFRAME Model & Vocabulary 2.0](#)  
This page provides access to three available vocabulary views of the BIBFRAME Vocabulary. The vocabulary is comprised of RDF properties, classes, and relationships between and among them RDF properties, classes, and relationships between and among them.
- [BIBFRAME Pilot \(Phase One—Sept. 8, 2015 – March 31, 2016\): Report and Assessment](#)  
Acquisitions & Bibliographic Access Directorate, Library of Congress. (2016, June 16). This document describes Phase One of the Library of Congress' pilot to test the efficacy of BIBFRAME. It includes descriptions of the planning process, what was being tested, the results, and lessons learned that will assist the Library of Congress as it moves to Phase Two of assessing the BIBFRAME model and vocabulary.
- [Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services](#)  
Eric Miller, Eric, Ogbuji, Uche, Mueller, Victoria , and MacDougall, Kathy. (2012, Nov. 21). Library of Congress.  
This document provides an introduction and overview of the Library of Congress, Bibliographic Framework Initiative.
- [bibliotek-o: a BIBFRAME Ontology Extension](#)  
Bibliotek-o is an ontology extension which defines additions and modifications to BIBFRAME, intended as a supplement to the core BIBFRAME ontology. It provides a set of recommended fragments from external ontologies and an application profile based on its recommended models and patterns. Bibliotek-o ontology extension is a joint product of the Mellon Foundation-funded Linked Data for Libraries Labs and Linked Data for Production projects.
- [bib.schema.org](#)  
This is a bibliographic extension for schema.org. The page lists the types, properties, and enumeration values for use in describing bibliographic material using schema.org.
- [DataCite Metadata Schema](#)  
DataCite. (2019, August 16). The DataCite Metadata Schema provides a list of core metadata properties chosen for accurate and consistent identification of resources for citation and retrieval purposes. Recommended use instructions are provided.
- [Direct Mapping of Relational Data to RDF](#)  
Arenas, Marcelo, Bertails, Alexandre, Prud'hommeaux, Eric, Sequeda, Juan (editors). (2012 Sept.27).  
This document defines a direct mapping from relational data to RDF with provisions for extension points for refinements within and outside of the document.
- [FAST Linked Data](#)  
FAST (Faceted Application of Subject Terminology) is an enumerative, faceted subject heading schema derived from the Library of Congress Subject Headings (LCSH). The purpose of adapting the LCSH with a simplified syntax to create FAST is to retain the vocabulary of LCSH while making the schema easier to understand, control, apply, and use. The schema maintains upward compatibility with LCSH, and any valid set of LC subject headings can be converted to headings. The site provides access to searchFAST, a user interface that simplifies the process of heading selection, and to a Web interface for FAST Subject selection available at FAST.

- [JSON Schema](#)  
Version 7 (Draft). (2019 March 31). This is a vocabulary that provides for the annotation and validation of JSON documents. It can be used to describe data formats, provide human and machine-readable documentation, make any JSON format a hypermedia format, allow the use of URI templates with instance data, describe client data for use with links using JSON Schema., and recognize collection and collection items.
- [Metadata Authority Description Schema \(MADS\)](#)  
MADS is an XML schema for an authority element set used to provide metadata about agents (people, organizations), events, and terms (topics, geographics, genres, etc.). It serves to provide metadata about the authoritative entities used in MODS descriptions.
- [Metadata Object Description Schema \(MODS\)](#)  
MODS is a bibliographic element set that may be used for a variety of purposes, and particularly for library applications. MODS is an XML schema intended to be able to carry selected data from existing MARC 21 records as well as to enable the creation of original resource description records. It includes a subset of MARC fields and uses language-based tags rather than numeric ones, in some cases regrouping elements from the MARC 21 bibliographic format. It is maintained by the Library of Congress.
- [Metadata Object Description Schema \(MODS\) - Conversions](#)  
Access MODS mapping including MARC to MODS, MODS to MARC, RDA to MODS, Dublin Core (simple) to MODS, and MODS to Dublin Core (simple). Style sheets are also available on this page.
- [Music Encoding Initiative \(MEI\)](#)  
MEI is an XML DTD for the representation and exchange of comprehensive music information. MEI is a schema that provides ways to encode data from all the separate domains: logical; visual; gestural (performance); and analytical, commonly associated with music. It accommodates bibliographic description required for archival uses. It also addresses relationships between elements, cooperative creation and editing of music markup, navigation within the music structure as well as to external multimedia entities, the inclusion of custom symbols, etc. MEI can record the scholarly textual apparatus frequently found in modern editions of music.
- [R2RML: RDB to RDF Mapping Language](#)  
Das, Souripriya, Sundara, Seema, Cyganiak, Richard (editors). (2012, Sept. 27). This document describes R2RML, a language for expressing customized mappings from relational databases to RDF datasets. The mappings provide the ability to view existing relational data in the RDF data model, expressed in a structure and target vocabulary of the mapping author's choice.
- [R2RML: RDB to RDF Mapping Language Schema](#)  
This document defines the R2RML: RDB to RDF Mapping Language schema which is used to specify a mapping of relational data to RDF.
- [Schema.org](#)  
Schema.org is a vocabulary that can be used with many different encodings, including RDFa, Microdata and JSON-LD to mark up web pages and e-mail messages. Sponsored by Google, Microsoft, Yahoo and Yandex, the vocabularies are developed by an open community process which includes an extension mechanism to enhance the core vocabulary for specific knowledge domains. It's primary function is to provide web page publishers a means by which they can enhance HTML pages so they can be crawled by semantic search engines linking the pages to the web of data.

- [Text Encoding Initiative \(TEI\)](#)  
The Text Encoding Initiative (TEI) is a global consortium which develops and maintains a set of Guidelines which specify encoding methods for machine-readable texts. TEI Guidelines have been widely used by libraries, museums, publishers, and individual scholars to present texts chiefly in the humanities, social sciences and linguistics. The site provides information on resources, projects using TEI, a bibliography of TEI-related publications, and TEI related software including Roma, a web-based application to generate P5-compatible schemas and documentation, and OxGarage, a tool for converting to and from TEI. In addition, the site links to a page of tools for use with TEI resources.
- [Thema](#)  
Thema is a multilingual subject category schema designed for the commercial global book trade industry to meet the needs of publishers, retailers, trade intermediaries, and libraries. Thema aims to reduce the duplication of effort required by the many distinct national subject schema, and to eliminate the need for scheme-to-scheme mapping that inevitably degrades the accuracy of classification, by providing a single scheme for broad international use. It can be used alongside existing national schema.
- [XML 1.0](#)  
Bray, Tim, Jean Paoli, Jean, C. M. Sperberg-McQueen, Maler, Eve, Yergeau, François, eds. (2013, Feb. 7). XML 1.0 is a version of the Extensible Markup Language used to store and transport data on the Web. It is both human and machine readable.
- [XML Path Language \(XPath\) 2.0](#)  
Berglund, Anders, Boag, Scott, Chamberlin, Don, Fernández, Mary F., Kay, Michael, Robie, Jonathan, Siméon, Jérôme, (eds.) (2015, Sept. 7). 2nd edition. XPath is an expression language that uses a path notation for navigating through the hierarchical structure of XML documents. XPath 2.0 is a superset of XPath 1.0. It supports a richer set of data types and takes advantage of the type information that becomes available when documents are validated using XML Schema.
- [XQuery 1.0: An XML Query Language](#)  
Boag, Scott, Chamberlin, Don, Fernández, Mary F., Florescu, Daniela, Robie, Jonathan, Siméon, Jérôme, (eds.). (2015, Sept. 7). 2nd edition. This is a version of XQuery, A query language that uses the structure of XML to express queries across all kinds of data, whether physically stored in XML or viewed as XML via middleware.

## Dublin Core

---

- [Dublin Core Metadata Initiative](#)  
This site provides specification of all Dublin Core vocabulary metadata terms maintained by the Dublin Core Metadata Initiative, including properties, vocabulary encoding schemes, syntax encoding schemes, and classes.
- [DCMI Application Profile Vocabulary](#)  
Coyle, Karen, editor (2021, April 9). This vocabulary supports the specification of Dublin Core Tabular Application Profiles (DC TAP). It is used to create a table or spreadsheet that defines the elements of an application profile. The vocabulary is also provided as a comma separated value template for use in a tabular form.
- [DC Tabular Application Profiles \(DC TAP\) - Primer](#)  
Coyle, Karen, editor. (2021, April 3). This primer describes DC TAP, a vocabulary and a format for creating table-based application profiles.



- [dctap](#)  
DCMI. (2021). dctap is a module and command-line utility for reading and interpreting CSV files formatted according to the DC Tabular Application Profiles (DCTAP) model. This document explains the project, installation, sub-commands, model, configuration, and provides a glossary.
- [dctap-python](#)  
DCMI. dctap requires Python 3.7 or higher. This GitHub page provides information and documentation on installing tap-python.
- [dctap/TAPtemplate.csv](#)  
Coyle, Karen. (2020, December 2). Access the TAP csv template from this GitHub page.

## Legal Schemas

---

- [Akoma Ntoso](#)  
Akoma Ntoso is an initiative to develop a number of connected XML standards, languages and guidelines for parliamentary, legislative and judiciary documents, and specifically to define a common document format, a model for document interchange, data schema, metadata schema and ontology, and schema for citation and cross referencing.
- [Legislative Documents in XML at the United States House of Representatives](#)  
U.S. House of Representatives. This page provides Document Type Definitions (DTD) for use in the creation of legislative documents using XML, links to DTDs, and background information regarding legislative XML. Access element descriptions and content models for bills, resolutions, Amendments, and roll call votes. This initiative was conducted under the direction of the Senate Committee on Rules and Administration and the House Committee on Administration, and with the involvement of the Secretary of the Senate and the Clerk of the House, the Congressional Research Service, the Library of Congress, and the Government Publishing Office.
- [Electronic Court Filing Version 4.01 Plus Errata 01](#)  
OASIS. Angione, Adam and Cabral, James, editors. (2014, July 14). This specification defines a technical architecture and a set of components, operations and message structures for an electronic court filing system, and sets forth rules governing its implementation. It was developed by the OASIS LegalXML Electronic Court Filing Technical Committee.

## RELATED RESOURCES

---

- [Akoma Ntoso an open document standard for Parliaments](#)  
Palmirani, Monica, and Vitali, Fabio. (2014). World e-Parliament Conference. This set of slides describes an open XML standard for legal documents used in Parliamentary processes and judgments.
- [BIBFLOW: A Roadmap for Library Linked Data Transition](#)  
Smith, MacKenzie, Stahmer, Carl G., Li, Xiaoli, and Gonzalez, Gloria. (2017, March 14). University of California, Davis and Zepheira, Inc. This is the report of the BIBFLOW project which provides a roadmap for libraries to use to transition into Linked Data environment. Recommendations for a phased transition are provided, as well as an analysis of transition tools, workflow transitions, estimated training, and work effort requirements.
- [BIBFRAME Comparison Tool](#)  
This tool provides for the side-by-side conversion of MARCXML records from the Library of Congress database to BIBFRAME2 using a LCCN or record number. Records can be serialized in Turtle or RDF XML.

- [Hello BIBFRAME2.0: Changes from 1.0 and Possible Directions for the Future](#)  
Kroeger, Angela. J. (2016, October 20). Criss Library Faculty Proceedings & Presentations. 65. This presentation introduces the basics and history of the BIBFRAME model, and its relationship to RDF, FRBR, and RDA. It covers core classes, editors, mixing metadata, holdings, approaches, PREMIS, changes from BIBFRAME1.0, and more.
- [Library of Congress BIBFRAME Manual](#)  
Library of Congress. (Revised 2020, May). This is the training manual for the BIBFRAME Editor and BIBFRAME Database.

## Vocabularies

## Vocabularies

---

- [Artists' Books Thesaurus](#)  
This controlled vocabulary is for artists' books. The Thesaurus is administered by the Arts Libraries Society of North America (ARLIS/NA). The platform, currently in draft form, will offer an illustrated, user-friendly guide to exploring and finding vocabulary terms.
- [DCAT \(Data Catalog Vocabulary\)](#)  
DCAT is an RDF vocabulary designed to facilitate interoperability between data catalogs published on the Web. This document defines the schema and provides examples for its use.
- [Data Documentation Initiative \(DDI\)](#)  
DDI Alliance. (2021). DDI is a free international standard for describing data produced by surveys and other observational methods in the social, behavioral, economic, and health sciences. It can be used to document and manage different stages in the research data lifecycle, such as conceptualization, collection, processing, distribution, discovery, and archiving.
- [DOAP \(Description of a Project\)](#)  
DOAP is an XML/RDF vocabulary to used to describe software projects, and in particular open source projects. This site hosts the DOAP wiki, and provides links to DOAP validators, generators, viewers, aggregators, and web sites using DOAP.
- [Expression of Core FRBR Concepts in RDF](#)  
This vocabulary is an expression in RDF of the concepts and relations described in the IFLA report on the Functional Requirements for Bibliographic Records (FRBR). It includes RDF classes for the group 1, 2, and 3 entities described by the FRBR report and properties corresponding to the core relationships between those entities. Where possible, appropriate relationships with other vocabularies are included in order to place this vocabulary in the context of existing RDF work.
- [FOAF \(Friend of a Friend\) Vocabulary Specification](#)  
This specification describes the FOAF language used for linking people and information. FOAF integrates three kinds of network: social networks of human collaboration, friendship and association; representational networks that describe a simplified view of a cartoon universe in factual terms, and information networks that use Web-based linking to share independently published descriptions of this inter-connected world.

- [Language of Bindings \(LoB\)](#)  
Ligatus, University of the Arts London. Based on SKOS, LoB is a thesaurus which provides terms used to describe historical binding structures. LoB can be used as a lookup resource on the website or as a software service where terms can be retrieved through an application. It can also be used for learning about book structures and materials, the frequency of the occurrence of bookbinding components, or other aspects connected with the book trade.
- [Lexvo.org](#)  
Lexvo defines global IDs (URIs) for language-related objects, and ensures that these identifiers are dereferenceable and highly interconnected as well as externally linked to a variety of resources on the Web. Data sources include the Ethnologue Language Codes database, Linguist List, Wikipedia, Wiktionary, WordNet 3.0, ISO 639-3 specification, ISO 639-5 specification, ISO 15924 specification, Unicode Common Locale Data Repository (CLDR), et. al. The site provides mappings between ISO 639 standards and corresponding Lexvo.org language identifiers and downloads of Lexvo datasets. Search over 7,000 language identifiers with names in many languages, links to script URIs (Latin and Cyrillic scripts, Indian Devanagari, the Korean Hangul system, etc.), geographic region URIs, etc.
- [OLAC video game genre terms \(olacvgt\)](#)  
Online Audiovisual Catalogers Network (OLAC). (2019). Guidelines for OLAC video game genre terms (olacvgt). This vocabulary provides a list of video game genre terms, each of which has a corresponding MARC authority record. Links to the MARC authority records are provided.
- [PeriodO](#)  
Rabinowitz, Adam T., Shaw, Ryan, and Golden, Patrick. PeriodO is a period gazetteer which documents definitions of historical period names. Definitions include a period name, temporal bounds on the period, an implicit or explicit association with a geographical region, and must have been formally or informally published in a citable source. Period definitions are modeled as SKOS concepts. Temporal extent is modeled using the OWL-Time ontology.
- [Rights Statements](#)  
The Rights Statements vocabulary provides rights statements for three categories of rights statements - In Copyright, No Copyright, and Other. Statements are meant to be used by cultural heritage institutions to communicate the copyright and re-use status of digital objects to the public. They are not intended to be used by individuals to license their own creations. RightsStatements.org is a joint initiative of Europeana and the Digital Public Library of America (DPLA).
- [Texas Digital Library Descriptive Metadata Guidelines for Electronic Theses and Dissertations, Version 2.0](#)  
Potvin, Sarah, Thompson, Santi, Rivero, Monica, Long, Kara, Lyon, Colleen, Park, Kristi. These Guidelines, produced by the Texas Digital Library ETD Metadata Working Group, comprise two documents to guide and shape local metadata practices for describing electronic theses and dissertations. The Dictionary, which lays out the standard, and the Report lays out detailed explanations for rationale, process, findings, and recommendations.

### **Vocabulary of Interlinked Datasets (VOID)**

- [Describing Linked Datasets with the Void Vocabulary](#)  
Alexander, Keith, Cyganiak, Richard, Hausenblas, Michael, and Zhao, Jun. (2011, March 3). This document describes the Void model and how to provide general metadata about a dataset or linkset (and RDF triple whose subject and object are described in different datasets).

- [Vocabulary of Interlinked Datasets \(VoID\)](#)  
Digital Enterprise Research Institute, NUI Galway. (2011, March 6). This document describes the formal definition of RDF classes and properties for VoID, an RDF Schema vocabulary for expressing metadata about RDF datasets. It functions as a bridge between publishers and users of RDF data, with applications including data discovery, cataloging, and archiving of datasets.
- [WorldCat Linked Data Vocabulary](#)  
OCLC's WorldCat Linked Data uses a subset of terms from Schema.org/ as its core vocabulary. Access the list of classes, attributes, and extensions with this link.

## Notes

---

For the Getty Vocabularies, please see the Registries, Portals, and Authorities page.

## Wikibase and Wikidata

---

### Wikibase and Wikidata

---

Wikibase is the platform on which Wikidata, A Wikimedia Project, is built. It allows for multi-language instances. For Wikibase Use Cases, see the Wikibase Use Case box on the bottom of the Use Cases page.

## Wikimedia Movement

---

Wikimedia is a global movement that seeks to bring free education to the world via websites known as Wikimedia Projects. Wikimedia Projects are hosted by the Wikimedia Foundation. Some of these Projects are listed below. Access the full family of Wikimedia Projects [here](#).

- [WikiCite](#)  
Wikimedia. WikiCite. (2019, July 20). WikiCite is an initiative to develop a database of open citations and linked bibliographic data to better manage citations across Wikimedia projects and languages. Potential applications include ease of discovering publications on a given topic, profiling of authors and institutions, and visualizing knowledge sources in new ways.
- [Wikidata](#)  
Wikidata is a free linked database that acts as central storage for the structured data of Wikimedia projects including Wikipedia, Wikivoyage, Wikisource, and others. It can be read and edited by both humans and machines. The content of Wikidata is available under a free license, exported using standard formats, and can be interlinked to other open data sets on the linked data web.
- [Wikipedia](#)  
Wikipedia is the open source encyclopedia within the MediaWiki universe. A page in Wikipedia is an article to which Wikidata can link.
- [Wikimedia Commons](#)  
Wikimedia Commons is a repository of freely usable media files to which anyone can contribute. Media files from Wikimedia can be linked to Wikidata statements.
- [Wikiquote](#)  
Wikiquote is a free compendium of sourced quotations from notable people and creative works in every language and translations of non-English quotes. It links to Wikipedia for further information.

- [Wiktionary](#)  
Wiktionary. (2019, June 27). Wiktionary is the English-language collaborative Wikimedia Project to produce a free-content multilingual dictionary. It aims to describe all words of all languages using definitions and descriptions in English. It includes a thesaurus, a rhyme guide, phrase books, language statistics and extensive appendices.

## Wikibase

---

- [Wikibase](#)  
Wikibase was developed for Wikidata as an open source collection of applications and libraries for creating and sharing structured data as linked data entities and their relationships. It consists of a set of extensions to the MediaWiki software for storing and managing data (Wikibase Repository) and for embedding data on other wikis (Wikibase Client). Wikibase provides an editing interface for creating, updating, merging, and deleting item and property entities.
- [Wikibase DataModel](#)  
MediaWiki. (2019, May 19). This specification describes the structure of the data that is handled in Wikibase. It specifies which kind of information can be contributed to the system. The Wikibase data model provides an ontology for describing real world entities, and these descriptions are concrete models for real world entities. For a less technical explanation of the model, see the Wikibase DataModel Primer.
- [Wikibase DataModel Primer](#)  
MediaWiki. (2018, August 1). This primer gives a good introduction to the Wikibase data model. It provides an outline and an explanation of the different elements of the Wikibase knowledge base and describes the function of each.

### Wikibase Resources

- [Install Docker Compose](#)  
In order to install a usable instance of Docker Desktop, the installation package must contain Docker Compose. If an instance is missing Docker Compose, this page provides instructions for installing it.
- [Installing a stand-alone Wikibase and most used services](#)  
This GitHub page provides instructions for establishing a Wikibase instance. It was developed by a member of Wikimedia Deutschland e. V. and four other software developers.
- [Use cases for institutional Wikibase instances](#)  
Mendenhall, Timothy R., Chou, Charlene, Hahn, Jim, et.al. (2020, May). Developed informally by library staff at Columbia University, Harvard University, New York University, and the University of Pennsylvania this GitHub page provides a wealth of information for for institutions considering installing their own Wikibase instance. Covering a wide range of topics such as local vocabularies, authority control, organizational name changes, cross-platform discovery, multilingual discovery, pipeline to Wikidata and broader web discovery, digital humanities, database, metadata, and exploratory projects, and more, each topic also supplies a use case example.
- [Wikibase Consultants and Support Providers](#)  
Wikimedia. (2021, Jan. 14). This page list Wikibase service providers who may help with issues with Wikibase instances.

- [Wikibase Install Basic Tutorial](#)  
Miller, Matt. (2019, September). Semantic Lab at Pratt Institute. This tutorial provides instructions for setting up Wikibase using Docker. The tutorial uses Digital Ocean, and it requires setting up an account at Digitalocean.com.
- [Wikibase Roadmap 2020 High-level Overview \(Public\)](#)  
WikiMedia. (2021, Jan. 11). This is an interactive chart that describes Wikibase development initiatives, including Wikibase Features, Wikibase System Improvements, Partnerships & Programs, Documentation, Wikibase Strategy & Governance, and Developer Tools.

## Wikidata Alert

---

### [Wikidata:SPARQL query service/Blazegraph failure playbook](#)

Wikidata. (2021, Dec. 13). This Wikidata article describes the proposed steps the Wiki Media Foundation is considering in the event of a catastrophic failure of its SPARQL Query Service powered by Blazegraph. The failure would occur when the amount of query-able data in Blazegraph exceeds Blazegraph's limits.

## Wikidata

---

Wikidata is a free, collaborative, multilingual software application built from Wikibase components that can be read and edited by humans and machines. It collects structured data to provide support for Wikimedia Projects including Wikipedia, Wikimedia Commons, Wikivoyage, Wiktionary, Wikisource, and others. The content of Wikidata is available under a free license, exported using standard formats, and can be interlinked to other open data sets on the linked data web.

- [Wikidata Introduction](#)  
Wikidata. (2018, June 18). This page provides a quick overview of Wikidata, its function withing the Wikimedia Universe, and an introduction to Wikidata basics.
- [Wikidata List of Policies and Guidelines](#)  
Wikidata. (2019, January 16). This page lists the proposed and approved policies and guidelines that govern Wikidata covering a wide range of topics including Living people, Deletion policy, Sources, Editing restrictions, Statements, Sitelinks, Verifiability, Administrators, Property creators, CheckUser, and more.
- [Wikidata: Notability](#)  
Wikidata.Notability. (2019, October 10). This page describes the Wikidata policy that sets forth the criteria needed for an item to be acceptable in Wikidata. It provides a link to a list of Wikimedia pages that are not considered automatically acceptable in Wikidata, and a link to a list of items that have been considered acceptable, in accordance to the general guidelines on this page.
- [Wikidata: Property constraints portal](#)  
Wikidata. (2020, June 19). Help:Property constraints portal. This page provides information on property constraints including a list of types and links to pages explaining how the constraints should be applied.
- [Wikidata Sandbox](#)  
Wikidata. (2020, August 24). This page provides a link to the Wikidata Sandbox in which you can experiment and practice using Wikidata. For experimenting with creating new items and properties, use the test.wikidata link on this page.
- [Wikidata Tours](#)  
(2018, April 7). This page provides access to interactive tutorials showing how Wikidata works and how to edit and add data.

## Articles, Development Plans, & Reports

---

- [ARL White Paper on Wikidata Opportunities and Recommendations](#)  
Association of Research Libraries (ARL). (April, 2019). In Wikisource. This paper discusses joint efforts between ARL and Wikidata to explore a way to interlink Wikidata to sources of library data and provide libraries and other GLAM institutions the opportunity to get involved in contributing to modeling and data efforts on a larger scale. Some possible contributions include name authorities, institutional holdings, and faculty information. Suggestions for contributing to Wikidata are also explored.
- [Creating Library Linked Data with Wikibase: Lessons Learned from Project Passage](#)  
OCLC Research. (2019, August). This document describes OCLC's Project Passage, a Wikibase prototype in which librarians from 16 US institutions experimented with creating linked data to describe resources without requiring knowledge of the technical machinery of linked data. The report provides an overview of the context in which the prototype was developed, how the Wikibase platform was adapted for use by librarians, and discusses use cases where participants describe creating metadata for resources in various formats and languages using the Wikibase editing interface. The potential of linked data in library cataloging workflows, the gaps that must be addressed before machine-readable semantic data can be fully adopted and lessons learned are also addressed.
- [Differences between Wikipedia, Wikimedia, MediaWiki, and wiki](#)  
MediaWiki. (2019, April 19). This article provides a brief description of components and related software of the Wikimedia movement. It also provides links to additional Wikimedia movement resources.
- [Introducing Wikidata to the Linked Data Web](#)  
Erleben, Fredo, Gunter, Michael, Krotzsch, Markus, Mendez, Julian, and Vrandečić, Denny. (2014). This document explains the Wikidata model and discusses its RDF encoding. It is a good place to start if you are considering editing Wikidata.
- [Lexemes in Wikidata: 2020 Status](#)  
Nielsen, Finn Arup. (2020, May). Proceedings of the 7th Workshop on Linked Data in Linguistics, pages 82–86. This article discusses the use of lexemes in different languages, senses, forms, and etymology in Wikidata.
- [Wiki](#)  
Wikipedia. (2019, August 22). This Wikipedia article explains the features of a wiki knowledge base website and discusses the software, history, implementations, editing, trustworthiness, and other aspects of a wiki.
- [Wikidata:Development Plan \[2020\]](#)  
Wikidata. (2020, June 17). This page provides an interactive roadmap to the projects on which the Wikidata Development Team is working during 2020. Clicking on projects in the Wikidata matrix will provide information on projects under the categories: Increase Data Quality & Trust; Build Out the Ecosystem; Encourage More Data Use; Enable More Diverse Data and Users; and Other. The Wikibase matrix includes categories: Wikibase Features; Wikibase System Improvements; Partnerships & Programs; Documentation; Wikibase Strategy & Governance; and Developer Tools.

- [Wikidata: Development Plan \[2022\]](#)  
Wikidata. (2022, February 10). This page provides the roadmap for the Wikidata development team (Wikimedia Deutschland) for Wikidata and Wikibase for 2022. Highlights of the plan include empowering the community to increase data quality, strengthen underrepresented languages, increase re-use for increased impact, empowering knowledge curators to share their data, ecosystem enablement, and to connect data across technological & institutional barriers. Some objectives include better understanding of which organizations want to use Wikibase in the future and for what, ensure Wikibases can connect more deeply with each other and Wikidata to form an LOD web, user testing of federated properties in combination with local properties, and more.
- [Wikidata/Strategy/2019](#)  
Wikimedia. (2019, August 27). Wikidata/Strategy/2019. This page provides access to a product vision paper and three product strategy papers discussing possible future developments for Wikidata and Wikibase and a very ambitious role in shaping the future of the information commons through 2030. Topics discussed include strategies for making Wikimedia projects ready for the future; maintaining and supporting Wikimedia's growing content; ensuring the integrity of Wikimedia content; furthering knowledge equity; and enabling new ways of consuming and contributing knowledge. There is a strategy paper discussing Wikidata as a platform and another discussing the Wikibase ecosystem.
- [Wikimedia:LinkedOpenData/Strategy2021/Joint Vision](#)  
Wikimedia. Linked Open Data/Strategy 2021/Joint Vision. This document sets out the Wikibase and Wikidata joint vision for working in Linked Open Data. The document describes the vision, strategy, guiding principles, and approach to building out the Wikibase ecosystem.

## Wikidata Related Resources

---

- [Creating and editing libraries in Wikidata](#)  
Scott, Dan. (2018, February 18). Dan Scott's blog provides useful linked data information. This blog entry describes how to create a Wikidata item for a particular library. Properties useful for describing libraries and their collections are also provided.
- [DBpedia](#)  
DBpedia is a crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the Web. DBpedia provides the ability for sophisticated queries against Wikipedia, and to link the different data sets on the Web to Wikipedia data.
- [Linked Open Data Cloud Wikidata Page](#)  
This LOD Cloud page provides information about Wikidata, including download links, contact information, SPARQL endpoint, triples count, the Wikidata namespace, and more. It also provides examples of Wikidata concepts using information about Nelson Mandela in Wikidata.
- [MediaWiki](#)  
MediaWiki. (2019, June 14). This is the MediaWiki main page. MediaWiki is a multilingual, free and open, extensible, and customizable wiki software engine used for websites to collect, organize, and make available knowledge. It was developed for Wikipedia and other WikiMedia Projects. It includes an API for reading and writing data, and support for managing user permissions, page editing history, article discussions, and an index for unstructured text documents.



- [Practical Wikidata for Librarians](#)  
Wikidata. (2021, Feb. 11). Wikidata:WikiProject Linked Data for Production/Practical Wikidata for Librarians. This page provides a vast array of resources for librarians and archivists interested in editing Wikidata, and provides a space to share data models and best practices. Among the resources are instructional materials, policies, project recipes, verifiability, guidelines for describing entities in particular domains, constraint reports, user scripts, gadgets, and more.
- [User:HakanIST/EntitySchemaList](#)  
Wikidata. (2021, April 20). This is a list of schemas used for describing Entities in Wikidata compiled by Wikidata user HkanIST.
- [Wikidata editing with OpenRefine](#)  
Wikidata: Tools/OpenRefine/Editing. (2021, April 25). This page provides links to tutorials, videos and a reference manual demonstration how to use OpenRefine to add and edit items in Wikidata. It also demonstrates using MarkEdit with OpenRefine and Wikidata.
- [Wikidata in Brief](#)  
Wikimedia Commons. (2017, July 31). This document gives a one page overview of Wikidata.
- [Wikidata Query Service in Brief](#)  
Stinson, Alex. (2018, March). This document gives a one page overview of the Wikidata query service.
- [Wikimedia](#)  
Wikibooks. (2019, April 12). This open book provides information on how to use Wikis covering topics including editing, basic markup language, images, templates, categories, namespaces, administrative namespaces, user namespaces, Wikipedia, Wikimedia Commons, Wikibooks, Meta, Wikidata, Wikiversity, Wikispecies, Wikiquote, Wikivoyage, and more.
- [Works in Progress Webinar: Introduction to Wikidata for Librarians](#)  
OCLC Research. (2018, June, 12). This OCLC Webinara gives a brief introduction to Wikidata.

## WikiProjects

### WikiProject Universities

---

- [WikiProject Universities](#)  
Wikidata. (2019, August 21). The purpose of this WikiProject is to provide better coverage of universities and other research institutions in Wikidata. The goal is to create a comprehensive and rich catalog of institutions, with strong links to other entities in the academic ecosystem (researchers, publications, alumni, facilities, projects, libraries...). The scope of the project includes listing the recommended statements about universities and evaluating their coverage across Wikidata; Building showcase items to demonstrate what an university item should ideally look like; Linking between items about universities and their subunits; Linking items about people to items of the universities they are/were educated at, work(ed) at or were/ are otherwise affiliated with; and providing counts by type, country, etc. Subpages and participants are listed.

- [WikiProject University of California](#)  
Wikipedia. (2019, February 17). This Wikipedia article describes the WikiProject to improve Wikipedia's coverage of the University of California system, which encompasses University of California campuses, professional schools, facilities and biographies of major figures. The site provides links to WikiProjects for all of the University of California campuses and the UC System.
- [WikiProject Stanford Libraries](#)  
Stanford Wikidata Working Group. (2019, September 4). This is the page for a WikiProject for work done at Stanford Libraries to connect library data with Wikidata. The page provides useful links, references, and guides covering a wide range of topics including Description guidelines, Wikidata policies and guidelines, Quick reference guides, Property resources, Projects, and much more.

## WikiProjects

---

- [WikiProject Books](#)  
WikiProject Books is used to: define a set of properties to be used by book infoboxes, Commons books templates, and Wikisource; map and import relevant data currently spread in Commons, Wikipedia, and Wikisource; and establish methods to interact with this data from different projects. Based on the Functional Requirements for Bibliographic Records (FRBR) model, Wikimedia projects uses a two level model, "work", and replaces the "expression" and "manifestation" levels of the FRBR model into one "edition" level. Bibliographic properties and qualifiers are listed here.
- [WikiProject Heritage institutions](#)  
Wikidata. (2019, October 3). This project aims to create a comprehensive, high-quality database of archives, museums, libraries and similar institutions. While the main focus is on institutions that have curatorial care of a collection, the scope of the project includes related institutions, such as lending libraries, exhibition centers, zoos, and the like, to the extent that they are not covered by any other WikiProject. The project also serves to coordinate a range pf activities including the creation of an inventory of all existing public databases that contain data about heritage institutions, the implementation and maintenance of ontologies and multilingual thesauri relating to heritage institutions, the ingestion of data about heritage institutions into Wikidata, the inclusion of the data into Wikipedia and its sister projects, through Wikidata-powered infobox templates or lists, and more.
- [WikiProject Libraries](#)  
Wikidata. (2019, September 25). The aims of this Wiki Project is to define a structure for libraries and to create and improve the items about library. The page provides item identifiers for types of libraries.
- [WikiProject Linked Data for Production/Practical Wikidata for Librarians](#)  
Wikidata. (2020, August 25). This project seeks to gather and organize resources for librarians interested in editing Wikidata as well as to prevent duplicative work and provide stepping stones and guidance for librarians interested in working with Wikidata. Resources, links to gadgets and user scripts, information on data modeling, and project recipes are provided.
- [WikiProject Maps](#)  
Wikidata:WikiProject Maps. (2019: April 25). This Wikidata page provides access to geographic projects in Wikidata, possible properties to use for maps entered into Wikidata, a list of map types, and a link to maps on Wikidata.

- [WikiProject Medicine/National Network of Libraries of Medicine](#)  
Wikimedia. (2019, September 17). The goals of this Wikipedia project are to Improve the quality of Wikipedia medical related articles using authoritative mental health resources, raise visibility of NLM mental health resources, and promote Wikipedia as an outreach tool for engagement and open data. The project is centered on an edit-a-thon for October and November, 2019.
- [WikiProject Museums](#)  
Wikidata. (2018, May 22). This project aims to define properties for items related to museums and the rules of use for these properties (qualifiers, datatypes, ...) and to organize the creation and improving the quality of the elements. The page provides suggested properties to use with museum related entities, tools, and example queries.
- [WikiProject Source MetaData](#)  
Wikidata. WikiProject Source MetaData. (2019, August 11). WikiProject Source MetaData aims to: act as a hub for work in Wikidata involving citation data and bibliographic data as part of the broader WikiCite initiative; define a set of properties that can be used by citations, infoboxes, and Wikisource; map and import all relevant metadata that currently is spread across Commons, Wikipedia, and Wikisource; establish methods to interact with this metadata from different projects; create a large open bibliographic database within Wikidata; and reveal, build, and maintain community stakeholderhip for the inclusion and management of source metadata in Wikidata. This page provides information regarding ongoing imports and projects, and a very substantial list of metadata sub-pages belonging to this project.
- [WikiProject Periodicals](#)  
Wikidata: WikiProject Periodicals. (2019, June 15). This project aims to: define a set of properties from w:Template:Infobox\_journal and w:Template:Infobox\_magazine (and other languages), especially prior names with year ranges, and standard abbreviations; define a set of properties about periodical publishers, including learned societies; map and import 'Journals cited by Wikipedia; map and import all relevant data to the Wikipedia collection of journal articles at w:Category:Academic journal articles / w:Category:Magazine articles (and other languages), and link these items to the reason for their notability - e.g. the discovery that was made, or event it records; prepare for linking Wikisource collection of journal/magazine articles into Wikidata; map and import all other relevant data that currently is spread in Commons, Wikipedia, and Wikisource; and establish methods to interact with this data from different projects. This page provides lists of properties relevant to periodicals.

## Wikidata Properties

## Wikidata Properties

---

- [Citing sources](#)  
Wikidata. (2019, January 5). This is a list of properties appropriate for citing sources in Wikidata. The list includes such properties as place of publication, imported from Wikimedia project, publisher, author, stated in, chapter, described by source, quote, inferred from, archive date, etc.
- [Wikidata List of Properties](#)  
Wikidata. (2019, July 22). This page provides access to Wikidata properties by broad description topics. The page also lists tools for browsing properties in different languages, and a download option for all properties.

- [Wikidata property for items about people or organisations](#)  
Wikidata. (2019, February 7). This is a list of properties that can be used to describe people or organizations. It encompasses a very wide range such as head of state, flag, logo, movement, league, chief executive officer, headquarters location, record label, Queensland Australian Football Hall of Fame inductee ID, field of work, award received, etc.
- [Wikidata property for items about people or organisations/human/authority control](#)  
Wikidata. (2019, October 5). This is a list of Wikidata name authority control properties for writers, artists, architects, and organizations.
- [Wikidata property for items about works](#)  
Wikidata. (2019, February 11). This is a list of properties to describe works such as articles, books, manuscripts, authority control for works, plays, media items, musical works, algorithms, software, structures, comics, television programs, works of fiction, and films.

## **Wikidata/Wikimedia Tools**

### **Wikidata/Wikimedia Tools**

---

There are many tools developed for working with Wikidata, many which are listed on the Wikidata Tools page listed below. General tools that are helpful with editing and adding items to Wikidata are listed here.

- [Author Disambiguator](#)  
Wikidata:Tools/Author Disambiguator. (2020, October 2). Author Disambiguator is a tool for editing authors of works recorded in Wikidata, and is partially coordinated with the Scholia project that provides visual representations of scholarly literature based on what can be found in Wikidata. By converting author strings into links to author items a much richer analysis and tracing of relationships between researchers and their works, institutions, co-authors, etc. can be achieved. The tools ability to integrate with Scholia provides enhanced visual analysis.
- [Cradle](#)  
Manski, Magnus. Cradle is a tool for creating new Wikidata items using a form. A link to existing forms along with their descriptions is provided. It is also possible to compose an original form.
- [Docker Desktop](#)  
Docker Desktop is a MacOS and Windows application for building and sharing containerized applications and microservices and delivering them from your desktop. It enables the leveraging of certified images and templates in a choice of languages and tools. Docker Desktop uses the Google-developed open source orchestration system for automating the management, placement, scaling, and routing of containers.
- [EntiTree](#)  
Schibel, Martin. EntiTree generates dynamic, navigable tree diagrams of people, organizations, and events based on information drawn from several sources and linked to Wikipedia articles.
- [FindingGLAMs](#)  
This tool is a modified version of Monumental used to display information and multimedia about cultural heritage institutions gathered through Wikidata, Wikipedia and Wikimedia Commons. Search by name of institution or explore by geographic region. example item, or city.

- [Miraheze](#)  
Miraheze is non-profit MediaWiki hosting service created by John Lewis and Ferran Tufan. The service offers free MediaWiki hosting, compatible with VisualEditor and Flow.
- [Monumental](#)  
Marynowski, Paweł and LaPorte, Stephen. (2017). This tool displays information and multimedia about cultural heritage monuments gathered through Wikidata, Wikipedia and Wikimedia Commons. Explore by entering a name of a monument, geographical region, example monument, or city.
- [OpenRefine](#)  
Wikidata:Tools/OpenRefine. (July 17,2019). OpenRefine is a free data wrangling tool used to clean tabular data and connect it with knowledge bases, including Wikidata. This page provides recipes, instructions, and resources to tutorials.
- [osm-wikidata](#)  
Betts, Edward. Downloaded October 22, 2019. Use this tool to match Open Street Map (OSM) Entities with Wikidata Items. It uses the Wikidata SPARQL query service and the OSM Overpass and Nominatim APIs. Installation and configuration instructions are provided.
- [Scholia](#)  
Nielsen, Finn Arup, Mietchen, Daniel, et. al. (2020). Scholia is a service which uses the information in Wikidata to create visual scholarly profiles for topic, people, organizations, species, chemicals, etc. It can be used with the Author Disambiguator tool to generate bar graphs, bubbles charts. line graphs, scatter plots, etc.
- [Semantic MediaWiki](#)  
Krötzsch, Markus. (2020, Apr. 19). Semantic MediaWiki (SMW) is an open source extension for MediaWiki, the software that powers Wikipedia. It provides the ability to store data in wiki pages, and query it elsewhere, thus turning a wiki that uses it into a semantic wiki.
- [Wikidata:SourceMD](#)  
Wikidata. (2019, February 10). SourceMD, aka Source Metadata Tool, can be used to take the persistent Wikidata identifier for a scholarly article or book to automatically generate Wikidata items using metadata from scholarly publications. The tool works with these identifiers: ISBN-13 (P212); DOI (P356); ORCID iD (P496); PubMed ID (P698); and PMCID (P932).
- [Wikidata Tools](#)  
(2019, January 18). This page provides a list of tools to ease working with Wikidata, including a property list, query tools, lexicographical data tools, tools for editing items, data visualization tools, a Wikidata graph builder, and more.
- [Wikimedia Programs & Events Dashboard](#)  
Wikimedia. (2020, January 21). The Programs & Events Dashboard is a management tool used to initiate and organize edit-a-thons, campaigns, and other wiki events. It provides instructions, registration functions, tracking functions to measure and report the outcome of a program (number of editors, number of edits, items created, references added, number of views, etc.).

## **Workshops and Projects**

## **Workshops and Projects**

---

This page provides access to documents and reports associated with workshops, institutions, organizations, or other entities which relate valuable information, or describe initiatives or projects regarding the Semantic Web or Linked Data.

- [Addressing the Challenges with Organizational Identifiers and ISNI](#)  
Smith-Yoshimura, Karen, Wang, Jing, Gatenby, Janifer, Hearn, Stephen, Byrne, Kate. (2016). This webinar discusses documenting the challenges, use cases, and scenarios where the International Standard Name Identifier (ISNI) can be used to disambiguate organizations by using a unique, persistent and public URI associated with the organization that is resolvable globally over networks via specific protocols, thus providing the means to find and identify an organization accurately and to define the relationships among its sub-units and with other organizations.
- [BIBCO Mapping BSR to BIBFRAME 2.0 Group: Final Report to the PCC Oversight Group](#)  
BBIBCO Mapping BSR to BIBFRAME 2.0 Group. (2017, July). This report summarizes the BIBCO Mapping the BIBCO Standard Record (BSR) to BIBFRAME 2.0 group's work and identifies issues that require further discussion by the Program for Cooperative Cataloging (PCC).
- [BIBCO Standard Record to BIBFRAME 2.0 Mapping](#)  
BIBCO Mapping BSR to BIBFRAME 2.0 Group. (2017, July). This spreadsheet maps BIBCO Standard Record elements to BIBFRAME 2.0. Amid the information included in the spreadsheet are RDA instructions & elements, MARC coding, rda-rdf properties as defined in the RDA Registry, Triple statements needed to properly map the element, and specific instructions pertaining to elements.
- [BIBFLOW](#)  
BIBFLOW is a two-year project of the UC Davis University Library and Zepheira, funded by IMLS. Its official title is "Reinventing Cataloging: Models for the Future of Library Operations." BIBFLOW's focus is on developing a roadmap for migrating essential library technical services workflows to a BIBFRAME / LOD (LOD) ecosystem. This page collects the specific library workflows that BIBFLOW will test by developing systems to allow library staff to perform this work using LOD native tools and data stores. Interested stakeholders are invited to submit comments on the workflows developed and posted on this site. Information from comments will be used to adjust testing as the project progresses.
- [BIBFLOW: A Roadmap for Library Linked Data Transition](#)  
Smith, MacKenzie, Stahmer, Carl G., Li, Xiaoli, and Gonzalez, Gloria. (2017, March 14). University of California, Davis and Zepheira, Inc. This is the report of the BIBFLOW project which provides a roadmap for libraries to use to transition into Linked Data environment. Recommendations for a phased transition are provided, as well as an analysis of transition tools, workflow transitions, estimated training, and work effort requirements.
- [BIBFRAME \(Bibliographic Framework\) Initiative](#)  
This is the homepage for BIBFRAME the Library of Congress' Bibliographic Framework Initiative. BIBFRAME is a replacement for MARC and serves as a general model for expressing and connecting bibliographic data to the Web of Data. Access links to general information, the vocabulary, BIBFRAME implementation register, tools, draft specifications for Profiles, Authorities, and Relationships, a BIBFRAME testbed, webcasts and presentations, and more.
- [British Library Data Model](#)  
This is the British Library's data model for a resource.
- [British Library Data Model - Book](#)  
This is the British Library's data model for cataloging a book in a Semantic Web environment.

- [British Library Data Model - Serial](#)  
This is the British Library's data model for cataloging a serial in a Semantic Web environment. This is the British Library's data model for cataloging a serial in a Semantic Web environment.
- [Common Ground: Exploring Compatibilities Between the Linked Data Models of the Library of Congress and OCLC](#)  
Jean Godby, Carol and Denenberg, Ray. (2015, Jan.). Library of Congress and OCLC Research. This white paper compares and contrasts the Bibliographic Framework Initiative at the Library of Congress and OCLC's efforts to refine the technical infrastructure and data architecture for at-scale publication of linked data for library resources in the broader Web.
- [CONSER CSR to BIBFRAME Mapping Task Group: \[Final Report\] of the PCC BIBFRAME Task Group](#)  
CONSER CSR to BIBFRAME Mapping Task Group. (2017). This report summarizes the mapping outcomes and recommendations of the group for mapping CONSER Standard Record (CSR) elements to BIBFRAME 2.0. It also identifies several issues that will require further discussion.
- [CONSER Standard Record to BIBFRAME 2.0 Mapping](#)  
CONSER CSR to BIBFRAME Mapping Task Group. (2017, July). This spreadsheet maps the CONSER Standard Record (CSR) elements to BIBFRAME 2.0. The spreadsheet "Examples" column contains links to sample code documents containing Turtle serializations of each CSR element in BIBFRAME.
- [CUL Metadata Application Profiles](#)  
Cornell University Library Metadata Application Profiles. This page provides an overview and documentation of Cornell University Library's use of metadata application profiles (MAPs). The page offers a definition and explains the role of MAPs in an application or metadata service, and gives examples. A wealth of information regarding documentation for training, MAPs used at CUL, and the CUL metadata ecosystem is provided.
- [Europeana pro](#)  
Europeana Foundation. This site provides a detailed description of the European Union's Linked Open Data initiative, including a history, the Europeana Data Model, a list of namespaces used, tools, and more.
- [Game Metadata and Citation Project \(GAMECIP\)](#)  
This University of California Santa Cruz and Stanford University project is developing the metadata needs and citation practices surrounding computer games in institutional collections. It seeks to address the problems of cataloging and describing digital files, creating discovery metadata, and providing access tools associated with the stewardship of digital games software stored by repositories. The site provides information regarding tools and vocabularies under development.
- [IIIF Explorer](#)  
OCLC ResearchWorks. (2020). The IIIF Explorer is a prototype tool that searches across an index of all of the images in the CONTENTdm digital content management systems hosted by OCLC.
- [Library of Congress Labs](#)  
The Library of Congress Labs site shares experimental initiatives the Library is conducting with its digital collections. Access videos, reports, presentations, and APIs. Clicking on the LC for Robots tab provides bulk data for Congressional bills, MARC records (in UTF-8, MARC8, and XML), Chronicling America, and more. The site demonstrates how to interact with the Library's collection.
- [Linked Art](#)  
Linked Art is a Community working on creating a shared model to describe art based on Linked Open Data. The site lists partner projects, consortia, and institutions.

- [Linked Data for Libraries \(LD4L\)](#)  
LD4L is a collaborative project of Cornell University Library, the Harvard Library Innovation Lab, and the Stanford University Libraries. The project is developing a Linked Data model to capture the intellectual value added to information resources when they are described, annotated, organized, selected, and used, along with the social value evident from patterns of usage.
- [Linked Data for Production: Closing the Loop \(LD4P3\)](#)  
LD4P3 aims to create a working model of a complete cycle for library metadata creation, sharing, and reuse. LD4P3 builds on the foundational work of LD4P2: Pathway to Implementation, LD4P Phase 1, and Linked Data for Libraries Labs (LD4L Labs). Access the statement of objectives for two domain projects, one for cartographic material and one for film/moving image resources.
- [Linked Data for Production: Pathway to Implementation \(LD4P2\)](#)  
Futornick, Michelle. (2019, January 14). LD4P Phase 2 builds upon the work of Linked Data for Production (LD4P) Phase 1 and Linked Data for Libraries Labs (LD4L Labs). This phase marks the beginning of implementing the cataloging community's shift to linked data for the creation and manipulation of their metadata. Access information regarding the seven goals of Phase 2 outlined by the institutions collaborating on the project: Cornell; Harvard; Stanford; the University of Iowa School of Library and Information Science; and the Library of Congress and the Program for Cooperative Cataloging (PCC).
- [Linked Data Wikibase Prototype](#)  
OCLC Research. In partnership with several libraries, OCLC has developed a prototype to demonstrate the value of linked data for improving resource-description workflows in libraries. The service is built on the Wikibase platform to provide three services: a Reconciler to connect legacy bibliographic information with linked data entities; a Minter to create and edit new linked data entities; and a Relator to view, create, and edit relationships between entities.
- [Looking Inside the Library Knowledge Vault](#)  
Washburn, Bruce and Jeff Mixter, Jeff. (2015, Aug.26).  
This is a U-Tube recording of an OCLC Research Works in Progress webinar describing how OCLC Research is evaluating the Google Knowledge Vault model to test an approach to building a Library Knowledge Vault.
- [OCLC Data strategy and linked data](#)  
This page describes OCLC library bibliographic initiatives focusing on designing and implementing new approaches to re-envision, expose, and share library data as entities that are part of the Semantic Web.
- [RDA Input Form](#)  
The RDA Input Form is a proof-of-concept experiment created by the Cataloging and Metadata Services of the University of Washington to demonstrate that RDA cataloging (input) can be easily output in multiple schemas using a processing pipeline and mappings. The form focuses on PCC core and output is in RDA/RDF and BIBFRAME in RDF-XML. The experiment showed that output in these schemas can be generated in an automated fashion using a pipeline. Implications for future production cataloging systems is that input and output should not be directly tied to each other, and cataloging systems should have sufficient flexibility to output in multiple schemas, which can be achieved in an automated way.
- [Report of the Stanford Linked Data Workshop](#)  
This report includes a summary of the workshop agenda and a chart showing the use of Linked Data in cultural heritage venues for the workshop held at Stanford University June 27 - July 1, 2011.



- [rightsstatements.org](https://rightsstatements.org)  
This GitHub page provides access to the request for proposals issued by the International Rights Statements Working Group, a joint Digital Public Library of America (DPLA) and Europeana Foundation working group to develop and implement a technical infrastructure for a rights statements application, a content management system, and a server configuration, deployment, and maintenance implementation for rights management. Links to a PDF version of the request and a PDF version of the "Requirements for the Technical Infrastructure for Standardized International Rights Statements" are provided.
- [Schema Bib Extend Community Group](#)  
This W3C group was formed to discuss and prepare proposal(s) for extending Schema.org schemas for the improved representation of bibliographic information markup and sharing. Access the group wiki, contact information, a mailing list, information regarding joining the group, information about proposals, an RSS feed, and recipes and guidelines.
- [SHARE Virtual Discovery Environment project](#)  
Casalini Libri, @Cult, and participating libraries. The aim of this project (Share-VDE Project) is to design a flexible configuration that uses the paradigms of the Semantic Web to provide a way for libraries to handle their data related to information management, enrichment, entity identification, conversion, reconciliation, and publication processes of the Semantic Web as independently as possible. The project provides a prototype of a virtual discovery environment with a three BIBFRAME layer architecture (Person/Work, Instance, Item) established through the individual processes of analysis, enrichment, conversion, and publication of data from MARC21 to RDF. Records from libraries with different systems, habits, and cataloguing traditions were included in the prototype.
- [Stanford Linked Data Workshop Technical Plan](#)  
This report summarizes the output of the Linked Data in cultural heritage venues workshop held at Stanford University June 27 - July 1, 2011.
- [Stanford Tracer Bullets](#)  
Futornick, Michelle. (2008, August 6). This Stanford Linked Data production project focused on all the steps to transition to a linked data environment in four technical services workflows: copy cataloging through the Acquisitions Department, original cataloging, deposit of a single item into the Stanford Digital Repository, and deposit of a collection of resources into the Stanford Digital Repository.
- [Wikipedia + Libraries: Better Together](#)  
Wikipedia + Libraries: Better Together was an 18-month OCLC project to strengthen the ties between US public libraries and English Wikipedia which ended in May, 2018. Information provided includes how librarians use and contribute to Wikipedia, teach information literacies using Wikipedia, and use Wikipedia for events. Training materials are provided.

## Highlights

---

### [The Europeana Linked Open Data Pilot](#)

Haslhofer, Bernhard and Isaac, Antoine. Proc. In Int'l Conf. on Dublin Core and Metadata Applications 2011. This is the model developed to make metadata available from Europeana data providers as Linked Open Data. The paper describes the model and experiences gained with the Europeana Data Model (EDM), HTTP URI design, and RDF store performance.

## Use Cases

## Use Cases

---

This page provides links to examples of Linked Data currently in use.

- [BBC Academy: Linked Data](#)  
The British Broadcasting Company (BBC) is an early experimenter and adopter of Linked Data. The BBC Backstage project, working with Wikipedia, developed and produced content rich prototypes showing the potential of Linked Data. Explore this site to experience the hidden power seamless exploitation of Linked Data.
- [Becoming Data Native: How BIBFRAME Extensibility Delivers Libraries a Path to Scalable, Revolutionary Evolution](#)  
Miller, Eric. (2017). Zepheira and The Library.Link Network. This is a PowerPoint presentation by Eric Miller presented at the 2017 American Library Association conference. It describes how third party linked data library vendor Zepheira uses BIBFRAME in its iterations to connect library collections to the linked data cloud, including the Library of Congress collection.
- [BIBFRAME 2.0 Implementation Register](#)  
Library of Congress. The BIBFRAME 2.0 implementation register lists existing, developing, and planned implementations of BIBFRAME 2.0, the Library of Congress' replacement for MARC.
- [The British National Bibliography](#)  
The BNB is the single most comprehensive listing of UK titles, recording the publishing activity of the United Kingdom and the Republic of Ireland. It includes print publications since 1950 and electronic resources since 2003.
- [Dallas Public Library](#)  
This Dallas Public Library site demonstrates a Library.Link Network instance of library resources implemented by third party linked data vendor, Zepheira.
- [Data.gov](#)  
Data.gov is the open data initiative of the United States government. It provides federal, state and local data, tools, and resources to conduct research, build apps, design data visualizations, and more. Data are provided by hundreds of organizations and Federal agencies, and the code is open source. The data catalog is powered by CKAN, and the content seen is powered by WordPress.
- [data-hnm-hu - Hungarian National Museum Datasets](#)  
The Hungarian National Museum has made its Linked Data datasets available on datahub. As a means of familiarizing Hungarian librarians with BIBFRAME, the datasets were published so that the BIBFRAME and MARC descriptions were crossed linked. Conversion features work and entity recognition and name entities are linked to external datasets.
- [dblp computer science bibliography](#)  
Schloss Dagstuhl - Leibniz Center for Informatics. (2020, January 4). dblp is an on-line reference database providing free access to high-quality bibliographic meta-data and links to the electronic editions of computer science publications. When an external website that hosts an electronic edition of a research paper is known, a hyperlink together with the bibliographic meta-data is provided. Some links require subscriptions and some are open access.

- [Digital Public Library of America \(DPLA\)](#)  
The Digital Public Library of America is a portal that brings together and makes freely available digitized collections of America's libraries, archives, and museums. More than that, DPLA is a platform that provides developers, researchers, and others the ability to create tools and applications for learning, and discovery. This is a site worth exploring to see the next generation library. Click on Bookshelf to search for a book. Visit the Apps page to find ways of accessing DPLA's resources. DPLA uses Krikri, a Ruby on Rails engine for metadata aggregation, enhancement, and quality control as part of Heiðrún, its metadata ingestion system.
- [DTU Orbit - The Research Information System](#)  
DTU Orbit is the official research database of the Technical University of Denmark, DTU. Available to browse in standard web browsers and in addition to providing open access to articles, it provides a linked data type graph interface to cross search publications, projects, activities, department profiles and staff profiles related to publications to which DTU employees have contributed.
- [English Language Books listed in Printed Book Auction Catalogues from 17th Century Holland](#)  
Alexander, Keith. This datahub dataset lists books in the English language section of Dutch printed book auction catalogues of collections of scholars and religious ministers. For access to this data set and other auction catalogues, see the Printed Book Auction Catalogues resource.
- [Europeana Pro](#)  
Europeana Foundation. This is the European Union's initiative to share its countries' rich cultural heritage resources. Information regarding APIs, tools, grants, and events are also provided.
- [Harvard LibraryCloud APIs](#)  
Created by Licht, Jeffrey Louis, last modified by Wetherill, Julie M. (2019, May 6). Library Cloud is a metadata service that provides open, programmatic access to item and collection APIs that provide search access to Harvard Library collections metadata.
- [Ligatus](#)  
Ligatus. (2021). Ligatus is part of an initiative of the University of the Arts London conducting research on documentation in historical libraries and archives. Some of the projects include the Language of Bindings Thesaurus, Linked Conservation Data, Artivity (a tool capturing contextual data produced during the creative process of artists and designers while working on a computer), The St. Catherine's Project (conservation support for the unique monastery library in Sinai), and Archive as Event (online archive of the artist John Latham structured using Creative Archiving principles based on Latham's ideas).
- [Linked Jazz](#)  
This Pratt Institute project is built around oral histories of jazz musicians from Rutgers Institute for Jazz Studies Archives, Smithsonian Jazz Oral Histories, the Hamilton College Jazz Archive, UCLA's Central Avenue Sounds series, and the University of Michigan's Nathaniel C. Standifer Video Archive of Oral History. Tools developed for the project include the Linked Jazz Transcript Analyzer, a Name Mapping and Curator Tool, the crowd sourcing tool Linked Jazz 52nd Street, and the Linked Jazz Network Visualization Tool. The project also used Ecco! - a Linked Open Data application for entity resolution designed to disambiguate and reconcile named entities with URIs from authoritative sources.
- [London DataStore](#)  
The London DataStore is a free and open data-sharing portal providing access to over 500 datasets about London.

- [National Széchényi Library catalogue \(National Library of Hungary\)](#)  
The National Széchényi Library provides an example of a library Linked Data interface. Use the search box to perform a search. Click on "Semantic Web" under "Services" and click on Semantic web to learn more about this library's service and its move to Virtuoso.
- [OCLC Research](#)  
This page shows OCLC's current research projects on libraries, metadata, collections, library enterprises, and more.
- [Office of the Historian](#)  
Department of State, United States. The Office of the Historian publishes the Foreign Relations of the United States and a Guide to Country Recognition and Relations, and the World Wide Diplomatic Archives Index. Among other resources provided by the Office are bibliographic information about U. S. Presidents and Secretaries of State, information about travels of the President and Secretaries of State, Visits by Foreign Heads of State, and more. The office is using the TEI Processing Model and eXistdb for publishing its documents on the Web.
- [Organization Name Linked Data](#)  
The Organization Name Linked Data (ONLD) is based on the North Carolina State University Organization Name Authority, a tool maintained by the Libraries' Acquisitions & Discovery department to manage the variant forms of name for journal and e-resource publishers, providers, and vendors in their local electronic resource management system (ERMS). Names chosen as the authorized form reflect an acquisitions, rather than bibliographic, orientation. Data is represented as RDF triples using properties from the SKOS, RDF Schema, FOAF and OWL vocabularies. Links to descriptions of the organizations in other linked data sources, including the Virtual International Authority File, the Library of Congress Name Authority File, Dbpedia, Freebase, and International Standard Name Identifier (ISNI) are provided.
- [SHARE Catalogue](#)  
@ Cult Rome Italy. Scholarly Heritage and Access to Research Catalog (SHARE Catalogue) is a portal providing a single point of access to the entirety of the integrated resources of eight Italian libraries organized according to the BIBFRAME linked data model.
- [Share-VDE \(Virtual Discovery Environment\)](#)  
Share-VDE is a library-driven initiative which collects the bibliographic records and authority files in a shared discovery environment using Linked Data. It is a collaborative endeavor between Casilini Libri, @CULT, the Program for Cooperative Cataloging, international research libraries, and the LD4P project. The Share-VDE interface provides wide-ranging and detailed search results to library patrons. Each library received the information corresponding to its own catalog in Linked Data which may be re-used according to local requirements with no restrictions.
- [Text Creation Partnership \(TCP\)](#)  
The TCP is making available standardized, accurate XML/SGML encoded electronic text editions from Early English Books Online (EEBO-TCP), Eighteenth Century Collections Online (ECCO-TCP), Evans Early American Imprints (Evans-TCP), and EEBO-TCP Collections: Navigations. Texts are from ProQuest's Early English Books Online, Gale Cengage's Eighteenth Century Collections Online, and Readex's Evans Early American Imprints and are made available through through web interfaces provided by the libraries at the University of Michigan and University of Oxford.

- [University of Edinburgh Wikimedian in Residence](#)  
University of Edinburgh. (2021). This page lists Wikidata Use Cases from the University of Edinburgh's collaboration with Wikimedia UK. Cases which have garnered international acclaim and served as inspiration for other research and collaborations include Scottish Witches, The Aberdeen Tower Block Archives, Documenting Biomedical Sciences: The Gene Wiki Project, Mapping the Scottish Reformation, Digitising Collections at the National Library of Wales, and others. Projects developed student skills as they surfaced data from MS Access databases to Wikidata as structured, machine-readable, linked open data.
- [University of Southampton Open Data Service](#)  
University of Southampton Open Data service has developed several mobile apps based on datasets using linked data. The data sets cover all aspects of university life including academic sessions, campus map, buildings, disabilities informaton, food services, organizations, and more. This initiative won the Times Higher Award in 2012 for Outstanding ICT Initiative of the Year, and a Cost Sector Catering award in 2015 for Best Innovation in Catering.

## Wikibase Use Cases

---

- [Enslaved](#)  
Michigan State University, Matrix: Center for Digital Humanities & Social Sciences. This Wikibase instance provides for the exploration of individuals who were enslaved, owned slaves, or participated in the historical trade. Search over numerous datasets and browse interconnected data, generate visualizations, and explore short biographies of enslaved and freed peoples.
- [The EU Knowledge Graph](#)  
European Commission. (2021, March 29). This Wikibase instance contains structured information about the European Union. Click on the Kohesio link to see the Project Information Portal for Regional Policy, which showcases how linked data can be uses to provide local policy information regarding different topics.
- [Rhizome Artbase](#)  
Rhizome provides a dataset for born-digital artworks from 1999 to the present day using the Wikibase platform. Search by date or artist name. Some entries include external links to artworks maintained by artists or others, archived copies hosted on Rhizome infrastructure, and documentation. The instance provides timeline capability and uses its own ontology data model that integrates with Wikidata and other standards.

## University of Edinburgh Wikimedian in Residence Projects

---

- [Mapping the Scottish Reformation](#)  
This project maps the Scottish Reformation by tracing clerics across early modern and modern Scotland using information from a database of the Scottish clergy generated by Wikidata. Information from this database runs parallel to another University of Edinburgh project, Scottish Witches.
- [Sottish Witches](#)  
Access the data visualizations of geolocation information pulled from the Survey of Scottish Witchcraft by Geology and Physical Geography student Emma Carroll. The work transformed the Survey from a static database to an acclaimed interactive linked open data collaboration with Wikimedia and with the support from Ewan McAndrew, University of Edinburgh's Wikimedian in Residence. Information about the project is available.

The mission of the UCLA Library is to provide access to and delivery of information resources to UCLA students, faculty, and staff in support of the research and instructional mission of the university. Relying on its highly skilled staff, the Library encourages innovation, capitalizes on appropriate technologies, forges effective partnerships and aggressively promotes excellence.