

# Be Empowered! Take the 'Scary' out of Linked Data

## Linked Data Essentials

**Linked Data** describes a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF –Wikipedia

The **Semantic Web** is an extension of the **World Wide Web**, the **Web of Data**. It is about making links between datasets that are understandable not only to humans, but also to machines. Linked Data is a set of best practices for sharing machine-readable interlinked data on the Web. It is sharing and using information, not just the data, but data put into context. Linked data shows great promise to enhance discovery and aids in revealing archival collections.

### Remember:

- Data needs to be machine readable.
- How? By creating data that can be used in Web architecture, ie Hypertext Transfer Protocol (**HTTP**) and Uniform Resource Identifier (**URI**).
- Data is expressed in triples. That means creating statements that are machine readable when transformed into a Resource Description Framework (**RDF**).
- Linked data can be queried using SPARQL (SPARQL Protocol and RDF Query Language) and other RDF query languages to retrieve and manipulate data stored in RDF.

## Linked Data Needs Two Components

### URI

*Uniform Resource Identifier (URI) is a recognized standard for identifiers*

- URIs used to uniquely identify anything.
- URIs play a key role in enabling Linked Data because they represent 'things' in a machine-readable form.
- URIs are used in HTTP web architecture.

### RDF

*Resource Description Framework (RDF) specifies relationships and transfers context with new data across applications*

- In other words: reusable data.
- RDF represents information in triples:
  - *Subject*: a resource, identified with a URI.
  - *Predicate*: specification of the relationship.
  - *Object*: a resource to which the subject is related.

### Principles of Linked Data

- Use URIs as names for things (people, organizations, artifacts, abstract concepts, etc.).
- Use HTTP URIs so that people can look up these names.
- When someone looks up a URI, offer valuable information using standards RDF and SPARQL.
- Go beyond describing the item by including links to other URIs so that users can discover other related items.

### Remember:

URIs for naming • RDF for describing • SPARQL for querying • Link to other URIs

Check out this link for more info!

<https://www.w3.org/DesignIssues/LinkedData.html>

## Key Questions

- How does our digital repository, CONTENTdm, support authority control?
- What if CONTENTdm does not fully accommodate linked data and cannot store data in RDF?
- What if we do not have the resources to maintain our own linked data infrastructure?
- Can we still participate in linked data under these conditions? What can we do to prepare our data? How can we incorporate linked data elements into workflows?

## The Project

The archives and library community has implemented linked data in recent years. Linked data empowers archivists to connect local data to a global audience using common identifiers and standards. However, due to the high level of institutional requirements that projects typically need with the high barrier of time and resources, many archivists have difficulty incorporating linked data practices into their daily descriptive work.

The Dr. Jo Ann Rayfield Archives at Illinois State University's Milner Library received the opportunity to digitize a segment of the expansive Ken-Way Studio Photograph Collection. The collection encompasses 120 linear feet and documents the history of Bloomington-Normal and the Central Illinois region. The Ken-Way Studio Digital Collection is a detailed photographic history of the development of national businesses in the Midwest. Many of the images in the digital collection feature the local and regional Central Illinois business community. Several companies held national significance, including Prairie Farms, Standard Oil, Admiral Corporation, and Nestle Beich.

Because the collection was not physically processed, it allowed an opening to invest time to research and create original descriptive metadata, including linked data elements, to accompany individual images in the digital collection. To lower the barrier, Milner Library observed an incremental approach to contribute to the linked data environment and prepared descriptive metadata for future use/reuse.



### Fitting the Pieces Together!

Linked data *does not have to be difficult* or require complicated local technical infrastructure. At Milner Library, we take a gradual approach to making data available for use on the semantic web for digital collections. We incorporate authoritative vocabularies and URIs as part of our descriptive metadata production workflow. Wikidata provides a platform where archives and libraries can contribute, explore, experiment and further the movement. It offers a low-barrier, high-result method for creating and using linked data in archives and libraries. It makes data not only visible but reusable as linked data.

For linked data implementations to succeed, it requires collaboration and consensus building between archives, cataloging and digital initiatives. Practical and informative outcomes should be planned, while also allowing room for iterative steps, experimentation and exploration with different tools and workflows.

Angela Yon, Cataloging & Metadata Librarian  
April Anderson-Zorn, University Archivist

Illinois State University, Milner Library and the Dr. Jo Ann Rayfield Archives,

## Vocabularies and Description

### Staged Approach to Linked Data at ISU

- Use authoritative controlled vocabularies.
- Search/add vocabulary terms and linked data values with URIs.



The Getty Thesaurus of Geographic Names (TGN)®

- Create Library of Congress Name Authority Records when necessary.

The Library of Congress > Linked Data Service > LC Name Authority File (LCNAF)

### Berglund, Kenneth H., 1917-2003

[Ken-Way Studio was started in 1945 by Kenneth Harold Berglund, local photographic company, United Photo, to create the base for his photographic timeline for many of the buildings, businesses and people had three children: Thomas, Kenneth, and Elizabeth. Berglund was Navy as a photographer.] <http://tempest.lib.ilstu.edu/Archon/?p=cc>

URI(s)

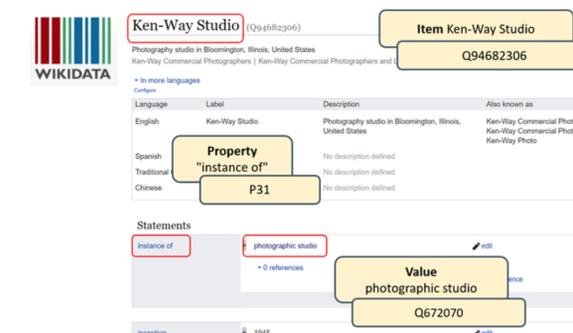
- <http://id.loc.gov/authorities/names/no2015045872>

The metadata production workflow focuses on high priority, low effort linked data tasks, such as adding URIs from the authoritative vocabularies. While the appropriate system and linked data technologies are not currently available at Milner now, the goal is to prepare data and have it ready for the foreseeable future.

## Wikidata

Wikidata is community-driven, like Wikipedia, and is a hub into free linked open data. It allows you to participate and make your data usable even without linked data technologies. Items can be created in just a few minutes.

Wikidata stores statements as triples = item + property + value



### BENEFITS!

- User-defined items & properties.
- Machine & human readable.
- Assigns a URI identifier at item creation - <https://www.wikidata.org/entity/Q#####>
- Creates RDF statements.

Wikidata offers the archives community a very simple way to engage with linked data, to contribute and share in the information ecosystem on Wikidata with GLAM (galleries, libraries, archives and museums) and other global organizations, individuals, and non-archives/library data. Wikidata now has 58,663,984 items. Simple searches are very quick and complex queries are supported by SPARQL. There is also the ability to experiment with queries once sharing your data in Wikidata with <http://query.wikidata.org>

## OAI and DPLA: Share the Data

### Open Archives Initiative - Protocol for Metadata Harvesting (OAI-PMH)

OAI is enabled in CONTENTdm for all of Milner Library's digital collections so data is harvestable by aggregators in other systems. The metadata can be harvested by others and reused/reconciled for implementation into linked data technologies.



Milner Library's digital collections are part of the regional DPLA hub, Illinois Digital Heritage Hub. As a member, others can export the metadata from your collections via Availability of Programmatic Interface (API). The development of API allows innovative reuse of data held in DPLA by developers and researchers both within and outside the network to build new knowledge.