UNLV | UNIVERSITY LIBRARIES

Library Faculty Presentations | Library Faculty/Staff Scholarship & Research

4-11-2013

Transforming digital collections into linked data: The rise of missing links

Silvia B. Southwick
University of Nevada, Las Vegas, silvia.southwick@unlv.edu

Cory K. Lampert
University of Nevada, Las Vegas, cory.lampert@unlv.edu

Follow this and additional works at: https://digitalscholarship.unlv.edu/libfacpresentation

Part of the Cataloging and Metadata Commons

Repository Citation

Southwick, S. B., Lampert, C. K. (2013, April). Transforming digital collections into linked data: The rise of missing links. Presentation at Association of College and Research Libraries National Meeting, Indianapolis, IN.

Available at: https://digitalscholarship.unlv.edu/libfacpresentation/109

This Poster is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Poster in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Poster has been accepted for inclusion in Library Faculty Presentations by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

Silvia Barcellos Southwick
Digital Collections Metadata Librarian
University of Nevada, Las Vegas
silvia.southwick@unlv.edu

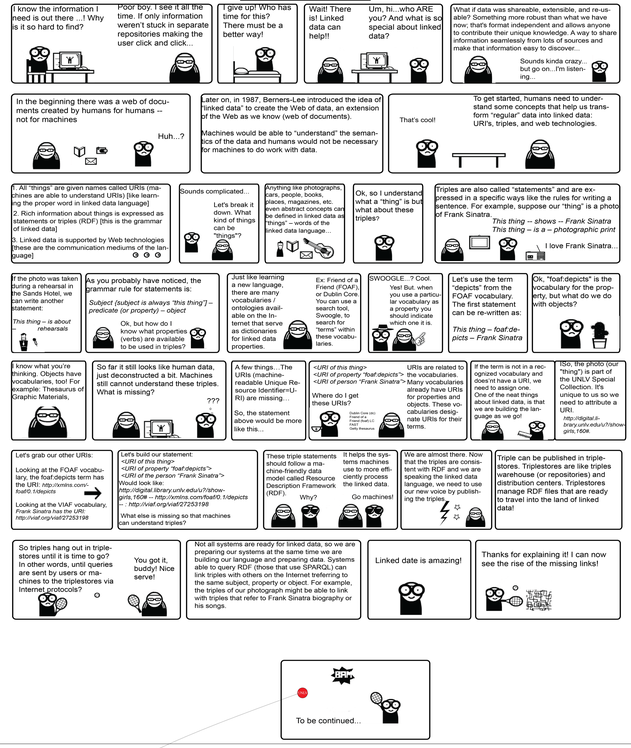
Transforming digital collections into linked data: The rise of missing links

Association of College and Research Libraries National Meeting

Indianapolis, IN April 11, 2013

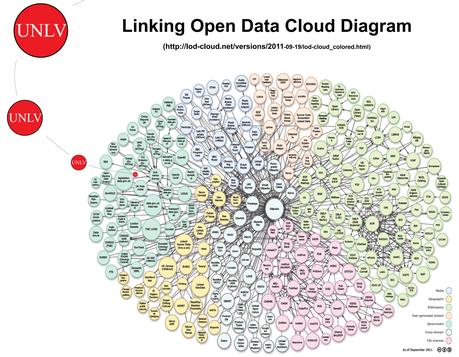
Cory Lampert
Head of Digital Collections
University of Nevada, Las Vegas
cory.lampert@unlv.edu

Rise Of the Missing Links

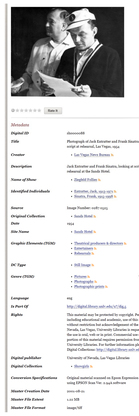


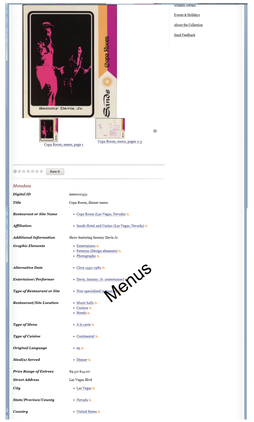
Linking Open Data Cloud Diagram

(http://lod-cloud.net/versions/2011-09-19/lod-cloud_colored.html)



Digital collection records => linked data graphical representation





Sample of possible triples:

this thing -- has title -- Photograph of Frank Sinatra and Jack Entratter reading a script at rehearsal, Las Vegas, 1954
this thing -- created by -- Las Vegas News Bureau
this thing -- has genre -- photographic print
this thing -- depicts -- Frank Sinatra
this thing -- depicts -- Jack Entratter
this thing -- created in -- 1954
this thing -- is about -- rehearsals
this thing -- is about -- entertainers
this thing -- has dc type -- image
this thing -- was taken at -- Sands Hotel
this thing -- has digital ID -- sho000088

Frank Sinatra -- has occupation -- entertainer
Jack Entratter -- has occupation -- theatrical producer

Sample of possible triples:

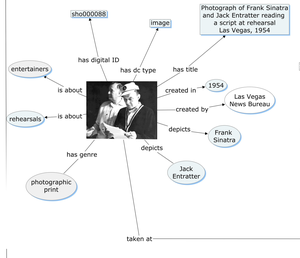
this thing -- has title -- Architectural drawing of Sands Hotel, additions and alterations, cocktail lounge, interior elevations, August 3, 1964
this thing -- created by -- Martin Stern Jr. A. I. A. Architect Associates
this thing -- has genre -- design drawings
this thing -- has drawing type -- interior elevations
this thing -- refers to component -- cocktail lounges
this thing -- was created in -- 1964-08-03
this thing -- is drawing of -- Sands Hotel
this thing -- has dc type -- image
this thing -- has digital ID -- sky000785

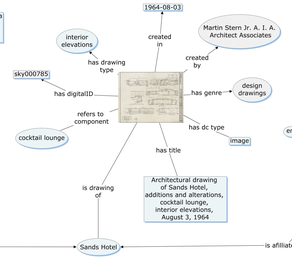
Sands Hotel -- organization type — resorts

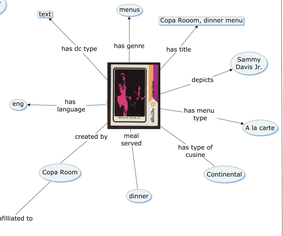
Sample of possible triples:

this thing -- has title -- Copa Room, dinner menu
this thing -- was created by -- Copa Room
this thing -- depicts -- Sammy Davis Jr.
this thing -- has genre -- menus
this thing -- has menu type -- A la carte
this thing -- type of cuisine -- Continental
this thing -- language -- eng
this thing -- meal served -- dinner
this thing -- has dc type -- text

Copa Room -- affiliated to -- Sands Hotel and Casino
Sammy Davis Jr. -- has occupation -- entertainers







UNLV Linked Data Project (exploratory study)

Guiding questions

- Is it feasible to create a common approach to transform metadata into linked data given that digital collections commonly adopt different metadata element sets?
- Is it possible to keep the richness and expressivity of our digital collection metadata using linked data?
- What model, if any, would be helpful in guiding this transformation?
- What available technologies would be useful to transform metadata into data?
- How will the generation of linked data impact current workflows for managing digital collections?

Goals

- Study the feasibility of developing a common process that would allow the conversion of our collection records into linked data *preserving their original expressivity and richness*
- Publish data from our collections in the Linked Open Data Cloud to improve discoverability and connections with other related data sets on the Web

How we got started

Formed a Library Linked Data Study Group to perform the following activities:

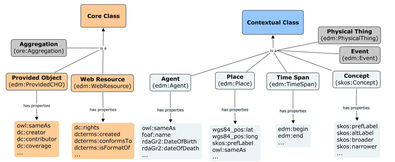
- Watch linked data webinars and discuss concepts
- Monitor literature in the linked data area and collect relevant documents
- Create an internal wiki with an extensive list of linked data resources
- Attend workshops and presentations on linked data
- Participate in linked data discussion lists (lita-l and BIBFRAME-l)

Designed the UNLV Linked Data Project to experiment with transforming digital collections records into linked data

- Formed a small project team
- Created a project proposal
- Created working groups charged to work on particular aspects of the metadata transformation such as: evaluating data models, identifying vocabularies for properties and objects, defining URIs for subjects, mapping metadata of our collections to a data model
- Created working group to research technological solutions
- Hired a developer/ GA to provide technological support to the implementation of the project

Europeana Data Model (EDM)

(http://pro.europeana.eu/edm-documentation)



Project phases

Preparation Phase

Evaluate, adopt or adapt technologies for:

- Converting records into linked data
- Creating / Managing triplestores
- Implementing query language (SPARQL)

Select *predicate* terms to be used for converting digital collection records into triples

Use the Application Profile metadata elements as basis for selecting properties.
Map Application Profile to the Europeana Data Model (EDM)

Attribute URIs to *subjects* and *objects*

Define rules for creating URIs / namespaces
If the "thing" to be identified with URI is "ours", attribute a new URI
If the "thing" has already an URI, use the existing URI

Define criteria for selecting a sample of records from the UNLV digital collections

Implementation Phase

Technology being tested

Export records	
Clean metadata	*OpenRefine*
Assign URIs	*OpenRefine/Manual*
Transform data	*OpenRefine*
Add external links	*OpenRefine/Manual*
Publish	*Mulgara server(testing) / SPARQL query language*

Assessment Phase

Evaluate benefits and challenges of transforming records into linked data:

- Comparing query results from traditional records vs. linked data
- Assessing technology solutions
- Analyzing impacts on workflow and metadata design

Lessons learned so far

Strategic lessons

- It is essential that the library as a whole, and administrators in particular, understand what linked data is and its potential impact on library systems
- Linked data concepts are new and unfamiliar in library environments where we are accustomed to work with "records" and MARC format. We have been exploring a variety of teaching methods to convey these concepts by testing various presentation of our project
- The project team should include professionals from various areas of the library as a way of raising awareness, disseminating linked data concepts, and preparing the library for possible future adoption of linked data
- It seems important to expose the project to external audiences to stimulate wider conversation

Technical lessons

- Our current system is not prepared to manage linked data
- The Europeana Data Model seems the best fit for guiding the transformation of our digital collections
- OpenRefine software is a flexible and powerful tool to support transformation
- The transformation of the digital collection records requires several metadata preparation steps