INTRODUCING NAVIGATOR: AN LOD INTERFACE FOR DIGITAL ASSETS

CORY LAMPERT
ARLIS 2017, NEW ORLEANS, LA
FEBRUARY 7, 2017
OVERVIEW

• The UNLV LOD Project
  – Feasibility of transformation of data
  – Data clean-up, preparation, workflows
  – Publication and visualization of the data

• Implications for users

• Tour of LOD technologies

• Navigator
  – Scope of project
  – Demo of the interface

• Questions
PT. 1: FOUNDATION OF THE UNLV LINKED DATA PROJECT (2014)

- Our digital collections consist of unique materials documenting the history of Southern Nevada stored in CONTENTdm; project focused initially on LOD for visual material collections

  • Definition of LOD we are using: “Linked Data refers to a set of best practices for publishing and interlinking data on the Web.”

  • A good way to better understand this is the 5-Star Data diagram: [http://5stardata.info/](http://5stardata.info/)
IN THE BEGINNING

GOAL:
Investigate the feasibility of a generalized process for transforming digital collections metadata into linked open data

Photo: Photograph of Howard Hughes in cockpit of the second XF-11, April 4, 1947, Howard Hughes Collection
**PREPARATION FOR LINKING**

**GOALS:**

Prepare metadata for linked open data by increasing the probability of linkages through quality.

Commit to the principle of preserving richness and empowering content contributors to describe materials.

Focus on controlled vocabularies to enhance data stream.

Photo: *Slide of a family in Boulder City, Nevada, circa early 1930s, Union Pacific Railroad Collections*
Most people want to see it!
This short video (no sound, just image) gives a preview of some existing applications to visualize what linked data may look like to users.

1. This video shows the Virtuoso Pivot Viewer software acting upon UNLV’s Linked Open Data – triplestore showing digital images from the Showgirls Collection.
   https://youtu.be/-83FTKEkYZ0

2. This video shows the Relfinder software acting upon UNLV’s Linked Open Data – triplestore showing graph-based relationships between oral history narrators in the Documenting African Americans in Las Vegas.
   https://youtu.be/wKCEI3KXdGk

Both of these applications have limitations, pros, and cons.
USER IMPACTS

GOAL:

Show linked data in action – beyond theory

Expose relationships between objects, agents, and collections for researchers

Utilize meaningful and compelling interfaces to browse/search the data in a scalable way

Photo: Five men with burros, circa 1900, Tonopah/Goldfield Collection
IMPLICATIONS FOR USERS

LOD AND DAMS
DAMS - NOW

A lot of results, no context

Shelley Berkley

Health care

Law

UNLV

Democrat

Student government

Mother

Yucca Mountain

Congresswoman

Israel
Linked data offers grammar for the mess of “words” traditional searching produces and provides meaningful results that communicate relationships between things in our repositories.

Start browsing and asking, “Who is Shelley Berkley?”

Shelley Berkley is the daughter of George Levine

Ask, “What did Shelley Berkley do?”

Shelley Berkley created transcripts of a roundtable
Shelley Berkley is referenced in a manuscript collection
Shelley Berkley is referenced in a photo collection
Shelley Berkley has the religion of Judaism
Shelley Berkley has successor Dina Titus

…and on and on….
DAMS - FUTURE

Moving from Digital Asset Management
to
Digital Assets with MEANING
QUICK TECHNICAL TOUR

SYSTEMS
STEPS IN DATA WORKFLOW
CURRENT TECHNOLOGICAL APPROACH

- CONTENTdm
- TemaTres
- OpenRefine
- OpenLink
- Virtuoso
- Names
- Metadata Management
- Controlled Vocabulary Management
- RDF files
- SPARQL Queries
- Triples
- Transformation To Linked Data
- Browser
CURRENT TECHNOLOGICAL APPROACH

Metadata Management

CONTENTdm

Names

TemaTres

Controlled Vocabulary Management

Metadata

OpenRefine

Transformation To Linked Data

RDF files

OpenLink Virtuoso

SPARQL Queries

Triple Store

Navigator

Browser

Triples
• Manage metadata (Create, Edit, Delete) allows:
  – Application Profile (creation)
  – Clean metadata (edit)
  – Share controlled vocabularies across collections (consistency across collections)

• Establish rigorous rules to use controlled terms
CURRENT TECHNOLOGICAL APPROACH

- CONTENTdm
  - Metadata Management

- TemaTres
  - Controlled Vocabulary Management

- OpenRefine
  - Transformation To Linked Data
  - RDF files

- OpenLink Virtuoso
  - Triple Store

- SPARQL Queries
  - Triples

- Browser

- Names
  - Metadata
  - Controlled Names
TEMATRES

• Open Source Package
  – Manages Controlled Vocabularies (workflow for approval of terms)
  – Allows to express relationships among terms

• Our Implementation
  – Person & Group CVs
  – Created specific Notes (text)
    • Biographical Notes
    • Resources that were consulted…
  – Recorded URIs from authority files
  – Added alternative names
  – Created new DB table to manage URIs
  – Adopted Relationships ontologies
  – Created Data Base Views to export data
CURRENT TECHNOLOGICAL APPROACH

Metadata Management
- CONTENTdm
- TemaTres

Metadata
- OpenRefine
- Local Controlled Names

Transformation To Linked Data
- RDF files
- SPARQL Queries

Triple Store
- OpenLink Virtuoso
- Triples

Controlled Vocabulary Management
- Browser

NAVIGATOR
• Open Source Package
  – Main page looks like a spreadsheet
  – Imports spreadsheets containing metadata
  – Generates RDF files with triples

• For the transformation of metadata into linked data
  – We created skeletons to map metadata to predicates (we adopted the Europeana Data Model)
CURRENT TECHNOLOGICAL APPROACH

- CONTENTdm
- TemaTres
- OpenRefine
- OpenLink
- Virtuoso

Metadata Management
Controlled Vocabulary Management

Metadata
Local Controlled Names

Names

Transformation To Linked Data

RDF files

SPARQL Queries

Triple Store

Browser
VIRTUOSO

- Triple store
- Uploads the RDF files generated by OpenRefine
- Triples may be associated to GRAPHS, which are identified by URIs
- At this stage of our implementation, each graph identifies the digital collection from which the data was generated. This might be revised in the future
- Data is currently published at: http://ld.library.unlv.edu/sparql
  - It is a SPARQL Endpoint
CURRENT TECHNOLOGICAL APPROACH

CONTENTdm

TemaTres

OpenRefine

OpenLink Virtuoso

Metadata Management

Controlled Vocabulary Management

Metadata

Local Controlled Names

Transformation To Linked Data

RDF files

SPARQL Queries

Triples

N A V I G A T O R

Browser
VISUALIZING THE DATA

DESIGN CONSIDERATIONS
Developed for the UNLV Jewish Heritage Project

It is a Browsing tool. Users can start by choosing terms from classes (People, Organizations, Collections, Videos, Audio, Photos)

Final phase of testing. It is installed in parallel to searching (not a replacement)

Developed by Alex Dolski (hired by UNLV for development)

Uses SPARQL queries to retrieve triples

Good tool to explore and discover relationships within the collection

At this point it has only data from UNLV collection but has potential to add data from DBPedia and possibly from other triple stores that might have related data
NAVI GATOR PANES

People Organizations Collections…

THIS THING

predicates

Subjects

Objects

predicates

predicates

Subject

Object
JHP NAVIGATOR

UNLV Jewish Heritage Project: http://digital.library.unlv.edu/jewishheritage

Experimental Navigator interface: http://lod.library.unlv.edu/nav/jhp


MEET YOU IN THE LINKED DATA CLOUD!
THANK YOU!

Research Team:
Cory Lampert: cory.lampert@unlv.edu
Silvia Southwick silvia.southwick@unlv.edu
Alex Dolski, developer
Meghan Gross, Digital Collections Specialist
Barbara Tabach, Southern NV Jewish Heritage Project Coordinator

Photo: Photograph of Bluebells posing outside of Pan Am jet, 1958, Donn Arden Collection