Library Data; From Linked Data to the Research Landscape

JOE MOCNIK, NORTH DAKOTA STATE UNIVERSITY
MACKENZIE SMITH, UNIVERSITY OF CALIFORNIA, DAVIS
Academic Research Landscape: A Setting for Developing Partnerships

JOE MOCNIK
NORTH DAKOTA STATE UNIVERSITY
There are challenges to the librarian and researcher since data sources are usually in silos and use different standards, rendering data integration difficult.

--2018 Top Trends in Academic Libraries
About NDSU

North Dakota State University is distinctive as a student-focused, land-grant, research university.
Overview
Most federal funding agencies now require a Data Management or Data Sharing Plan as part of a proposal submission. This is a formal document that outlines the type of data you are collecting, standards used to describe the data (metadata), who owns the data and how it can be accessed, considerations needed to protect sensitive information, including study participant confidentiality and intellectual property protection, and how you will ensure the archiving and preservation of the data. Carefully read proposal solicitations and agency guidelines for specific data plan instructions. Requirements may vary by agency and program.

For specific information on federal agency requirements, see Federal Agency Information and Open Access.

Data Life Cycle
Take into account the life cycle of data when developing data management plans. There are various stages to consider, including:

Proposal Planning and Project Start-up
- Collection
- Analysis
- Sharing
- Archiving
- Re-Using

Data management throughout the life-cycle of your research can be beneficial in numerous ways:
- Save time
- Increase research impact
- Ensure long-term ability to preserve fragile data sets
- Organize and categorize data for efficient access, analysis, queries, etc.
- Support sharing and open-access
- Focus on data sharing as an objective of investigation
- Support data-intensive discovery across disciplines
- Promote verification and replication of research analysis and findings
Opportunity

- Embrace data as valuable assets
- Uphold research in the digital age
- For investigators to conduct research
- For academic institutions to advance knowledge and develop competitive advantage
- For librarians to extend expertise to manage and preserve scholarship
Leverage diverse responsibilities

- Risk management
- Compliance & privacy
- Information technologies
- Systems administration
- Legal counsel
- Provost & other policy stewards

- Faculty researchers
- Faculty affairs
- Research management
- Library and archives
- Graduate education
Why librarians & research data?

Traditional roles

• Collecting
• Curating
• Preserving
• Describing
• Discovering
• Accessing
Why librarians & research data?

• Receptiveness for RDM
• Long term funding cycles
• Sustainable models to collect, preserve and access
• Trained info managers
• Partnerships with researchers
• Training next generation researchers
Create synergistic relationships across disciplines
Harness, integrate, and use intellectual energy
Take advantage of university knowledge assets
Empower faculty and students to become change agents
Borrow tools for leaning into the future of rapid technological change

- Missional shifting
- Transitional technologies
- Predicting the future
- Options for approaching the future
Evolve the organization beyond its hierarchy

- Embrace diverse partnerships to evolve the library
Intellectual entrepreneurship efforts

- Realign and train staff for new expertise
- Identify how researchers conduct work:
  - Use cases, observation, group reflections help understand challenges
  - Faculty survey identifies range of awareness
- Convene self-governance Forum of campus leaders
- Offer support services to improve productivity
Transforming Libraries for research support

- Mission to be a learning enterprise
- Changing metaphors about collections
- Transition beyond server administration
- Pragmatic combined view of the future
Evolve organization beyond its hierarchy

- Matrix direction
- Professional Forum
- Libraries Advisory Group
- Expand your skills…Expand your engagement
- Library Explorers
- Libraries Faculty Fellows
• Research output is an institutional asset
• Managing research output crosses multi-campus units
• Librarians empower change agents
• Partnering around research infrastructure is emerging
THANK YOU

Joe Mocnik
NORTH DAKOTA STATE UNIVERSITY
Josip.Mocnik@ndsu.edu
Pivoting from Prototype to Production: Making Linked Data in Libraries a Reality

MacKenzie Smith
University Librarian and
Vice Provost for Digital Scholarship
University of California, Davis

#oclcarc18
Linked Data at the UC Davis Library

2015-2016
- BIBFLOW

2017
- OCLC Prototype
- Local Triplestore

2018
- OCLC Prototype
- LD Browser
- LD4P

2019
- OCLD LD Production Cataloging
- LD Folio (Circulations and Acquisitions)
Sparse Local Database

- Store minimal amount of information possible
- Initially allowed two non-URI facets (probably going to change this)
- Rely on network and Linked Data ecosystem to deliver graphs with record equivalent information depth.
Local Data Ecosystem

- OCLC Prototype provides URI management and reconciliation.
- External data sources provide deep information context for delivering data to users.
- Mix of traditional and non-traditional data sources
Basic Cataloging
Entity Cataloging

- Add rich contextual information about entities
- Define relationships between entities
- Mint new entities
- Robust API allows machine traversal
Main Page

Welcome to Project Passage [edit]

With Project Passage, OCLC is partnering with a group of library resource-description workflows in libraries.

User Interface [edit]
http://18.218.102.193/entity/Q1225534

Xiaoli Li

Technical Services Librarian from University of California, Davis

<table>
<thead>
<tr>
<th>Language</th>
<th>Label</th>
<th>Description</th>
<th>Also known as</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Xiaoli Li</td>
<td>Technical Services Librarian from University of California, Davis</td>
<td></td>
</tr>
<tr>
<td>Chinese</td>
<td>No label defined</td>
<td>No description defined</td>
<td></td>
</tr>
</tbody>
</table>

Statements

Wikipedia (0 entries) edit

Wikinews (0 entries) edit

Wikiquote (0 entries) edit

Wikisource (0 entries) edit

Wikivoyage (0 entries) edit

Other sites (0 entries) edit
<table>
<thead>
<tr>
<th>WikiData</th>
<th>VIAF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sun Yat-sen</td>
<td>孫中山, 1866-1925</td>
</tr>
<tr>
<td>Chinese physician, politician and revolutionary</td>
<td>personal name</td>
</tr>
<tr>
<td>Dr Sun Yat-sen Museum</td>
<td>孫中山夫人，中華人民共和國名譽主席 宋慶齡</td>
</tr>
<tr>
<td>museum</td>
<td>personal name</td>
</tr>
<tr>
<td>Sun Yat-sen Museum Penang</td>
<td>孫中山記念會</td>
</tr>
<tr>
<td></td>
<td>corporate name</td>
</tr>
<tr>
<td>Sun Zhongshan</td>
<td>孫中山記念館</td>
</tr>
<tr>
<td>1986 film</td>
<td>corporate name</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>孫中山紀念館 (南京)</td>
<td>孫中山大元帥府紀念館</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fingerprint

Chinese

孫中山紀念館

Japanese

孫中山紀念館

Statements

country People's Republic of China

coordinate location 32°3'44N, 118°51'20E

"Same As" Identifiers

Library of Congress Name Authority File ID http://id.loc.gov/authorities/names/nr96028777

VIAF ID https://viaf.org/viaf/155392480

Wikidata Item ID http://www.wikidata.org/entity/Q10945437
Example: Mark Twain in Context

Connecting the UC Davis Catalog to the Semantic Web through the OCLC Pilot
The Science Fiction of Mark Twain
The Science Fiction of Mark Twain

Details

Title: The science fiction of Mark Twain
Author: Twain, Mark, 1835-1910.
Ketterer, David.

Related Persons: Ketterer, David.

Subjects: Science fiction, American

Description: Bibliography: p. [381]-385.

Publisher: Hamden, Conn. : Archon Books

Publication Date: 1984

Format: xxxiii, 385 p. ; 24 cm.

Identifier: ISBN : 0208020365 (alk. paper)

Language: English

Source: 01UCD_ALMA

System ID: 990013219810403126

Call number: PS1303 .K4 1984
The pilot’s managed Linked Data ecosystem of allows us to dramatically reduce the amount of data that must be locally stored and managed.
Networked Discovery

Step 1: User enters author name in search box based on VIAF Lookup.

Step 2: User submits search; Server pings Radorank for associated URLs.

Step 3: Server retrieves relevant collections of OCLC Work IDs associated with mentioned author (LDR) collection in local catalog.

Step 4: Server retrieves relevant collections of OCLC Work IDs and displays brief results to user.

Step 5: User selects item.

Step 6: User receives knowledge graph and holdings information.

User Interface

Server Application

Data Sources

Compile Collection of associated URLs

Search Local Catalog

Retrieve OCLC Works Graphs

Reconcile all LDRs in Work Graph

Build Knowledge Graph

VIAF

Reconciler

Local Implication

OCLC Works

Reconciler

Amazon API

ONIX API

Local Inventory System

Wiki Data
Rich Information Graph from Sparse Local Metadata

http://discover.library.ucdavis.edu
Detail of Contributor Information

http://discover.library.ucdavis.edu
Detail of Publication Information

http://discover.library.ucdavis.edu
LD Implementation

Next steps and future needs for production implementation

#oclcarc18
Hybrid Production Environment

OCLC Connexion
- Catalog as usual in Connexion
- Autofill entity labels and URIs
- Manage relationships in LD Prototype

Ex Libris Alma
- Send URIs in MARC

Ex Libris Developer Network
- Export LD to Local Database

Sparse Local LD Database

folio
future of libraries is open
Circulation & Acquisitions

#oclarc18
Future Ecosystem
Benefits of Project Participation

• Gained practical skills working with linked data
• Enhanced understanding of the value of linked data
• Helped imagine the future of cataloging
• Part of the linked data community, shaping how it will work
THANK YOU

Mackenzie Smith
UNIVERSITY OF CALIFORNIA, DAVIS
macsmith@ucdavis.edu