STARS

Faculty Scholarship and Creative Works

1-9-2016

Preparing for linked data in digital repositories

Sai Deng

University of Central Florida, sai.deng@ucf.edu

Original Citation

Deng, S. (2016). Preparing for linked data in digital repositories. ACRL Technical Services Interest Group Meeting. American Librarian Association (ALA) MidWinter Conference 2016. Boston, MA, January 9, 2016.

PREPARING FOR LINKED DATA IN DIGITAL REPOSITORIES

Sai Deng, University of Central Florida Libraries

How do digital repositories accommodate authority control?

Dspace

CONTENTdm

- Allows **controlled vocabulary** to be defined for each field;
- **Look up and add** terms;
- Need to either **define the controlled list yourself,** or **choose to use a built-in thesaurus** (from OCLC Terminologies Service).

- Fields can be configured to **allow authority control;**
- **Look up and add** terms;
- Need to define the controlled list, **keywords and subjects are not distinguished,** allow LCSH;
- Not linked to national or international authority database for control.

Digital Commons

- **Discipline thesaurus** functions as broad subjects;
- **Author Merge Tool:** merges various names but names are created based on author/submitter's email, not authority file;
- Allow fields to be defined as **link-friendly.**

Authority Control in Digital Repositories: a step towards a linked data world? A step away from the growing popular keyword search?

Authority Control in Digital Repositories and Related Linked Data Projects

■ **Myntti, J. & Cothran, University of Utah collaborated with Backstage Library Works.** Authority control in a Digital Repository: Preparing for Linked data. Adapted the process for MARC 21 authority control for digital collections @ CONTENTdm.

■ **Tarver et al., University of North Texas Libraries: UNT Names** (http://digital2.library.unt.edu/name/), link to other vocabularies such as VIAF, LC name authority file, Google Scholar author profile, LinkedIn profile.

■ Ideally, the systems can accommodate linked data or you have the resources to help you create external linked data and implement data reconciliation. If not, what can be done as a metadata librarian or a cataloging librarian?

The UCF Libraries' Experiment of Adding Linked Data Fields to Its IR @ Digital Commons

Creator (Linked data)	Cassanello, Robert (VIAF); Cassanello, Robert (LC)
Contributor (Linked data)	Doran, Glen H. (VIAF); Doran, Glen H. (LC); Milanich, Jerald T. (VIAF); Milanich, Jerald T. (LC); Florida Memory Project (VIAF); Florida Memory Project (LC); Florida. Division of Historical Resources (VIAF); Florida. Division of Historical Resources (LC); Brevard Museum of History and Natural Science (Cocoa, Fla.) (VIAF); Brevard Museum of History and Natural Science (Cocoa, Fla.) (LC)); Orange County (Fla.). Regional History Center (VIAF); Orange County Regional History Center (LC)
Collection	A History of Central Florida: Podcast
Geographic Coverage	Windover Pond Site (Fla.); Brevard County (Fla.)
Temporal Coverage	6000BC-5000BC
Contributor	Doran, Glen H.; Milanich, Jerald T.; Wentz, Rachel K.; Florida Memory Project; Florida. Division of Historical Resources; Brevard Museum of History and Natural Science (Cocoa, Fla.); Orange County Regional History Center; RICHES of Central Florida
Disciplines	Archaeological Anthropology	History of Art, Architecture, and Archaeology	Public History	Social History
Description	In this episode of A History of Central Florida we spoke with Drs. Glen Doran, Jerald Milanich, and Rachel K. Wentz about the Windover Burial Site available in replica form at The Brevard Museum of History and Natural Science.

Linked to VIAF or LC name authority uri(s)

Linked data fields:
o Use Digital Commons memo fields which allow links, e.g., Creator (Linked data), Contributor (Linked data);
o Both VIAF and LC linked data uri(s) are added for the names.

Statement: Subject – Predicate - Object
This podcast – has contributor – Doran, Glen H.

The UCF Libraries' Experiment of Adding Linked Data Fields to Its IR @ Digital Commons

The UCF Libraries' Experiment of Adding Linked Data Fields to Its IR @ Digital Commons

Title	Episode 01 Windover Burial Site
Series Producer	Robert Cassanello, *University of Central Florida*
Streaming Media	https://www.youtube.com/watch?v=-UYVBFHo-As
Media Format	embedly
Title - Alternative	Windover Burial Site
Preferred Title	A History of Central Florida: Podcast
Keywords	Archaic Indians; Archaic Florida Indians; Mortuary analysis; Prehistoric tools; Prehistoric bones; Prehistoric fabrics; Archaic period; Archaic society; Death; Land; Preparation in death; Bury ritual; Native Americans; Artifacts; Bury goods
Subjects	Windover Pond Site (Fla.); Paleo-Indians--Funeral customs and rites--Florida--Brevard County; Paleo-Indians--Anthropometry--Florida--Brevard County; Burial--Florida--Brevard County; Brevard County (Fla.)--Antiquities; Tools; Tools, Prehistoric--Analysis; Bone implements, Prehistoric; Textile fabrics, Prehistoric; Indians; Excavations (Archaeology)--Florida--Brevard County; Water-saturated sites (Archaeology); Archeology; Civilization
Type	Video; MovingImage

Keywords and Subjects:
- Separate fields;
- Some subjects have linked data uri(s), some subjects with subdivisions don't have individual uri(s);
- Both are only text fields; possible to add links;
- Digital Commons only display keywords as facets; need to request to add subjects as facets

Type, geographic coverage, etc.: can convert to linked data fields.

More possible statements:
This podcast – has subject - Indians
This podcast – has type – MovingImage
This podcast – has geographic coverage – Brevard County (Fla.)

Cataloging Digital Collections @ CONTENTdm: Adding Linked Data Values/Labels

- **CONTENTdm**: Adding **linked data values** to some fields for some collections
  - CONTENTdm doesn't "fully" accommodate linked data;
  - **Trying to catalog using linked data values for various fields,** such as: Name, Place, State, Country, subjects… (terms uri(s) not added)
  - Use "Subject" and "Keyword" fields separately;
  - Always **search LC Linked Data services, FAST and other online thesauri.**

- **Non-MARC data in digital repositories can take advantage of controlled vocabularies** used in traditional library catalogs and furthermore, linked data lists;

- Is it deemed good practice for the catalogers to **use linked data values/labels** to prepare for the linked data in digital repositories?

Some Reflections

■ **Limitations of Existing Authority Files:**

    – VIAF, LC Name Authority, Local Authority...   (Inadequate)

    – When automatic full validation and data reconciliation is not yet built-in to the systems, how much time can be spent checking and adding the linked data values?

    – Is it the ideal world: Keywords co-exist with controlled vocabularies, especially linked data lists? (More freedom vs. more control?)

■ **Reactions to Linked Data**

    – General initial reaction: Confusion, doubts;

    – Support experiments will try and evaluate...

Some Reflections

- **Collaboration and Consensus Building**
  - Collaboration between Cataloging and Digital Initiatives;
  - Collaboration between the library and other university units: The Regional Initiative for Collecting the History, Experiences, and Stories of Central Florida (RICHES), History Department...
  - Consensus building, experiments and evaluation

- **Possibilities:**
  - **Name lists in linked data form** (to meet local needs)
    - **Personal Names:** (link to) VIAF; LC authority name; Google Scholar profile; DBpedia; ORCID...
    - **Organizational Names:** (link to) VIAF; LC authority name headings...
    - **Geographic names, places:** (link to) LC authority geographic names; GeoNames...
  - **Create linked data from collections:** Create mapping between metadata elements and data model predicates; Implement reconciliation; Build interface for access and displaying of linked data

References and Readings

Lampert, Cory K. & Southwick, Silvia B. (2013). Transforming CONTENTdm metadata into linked open data. CONTENTdm Virtual Conference, December 2013.

Luyten, Bram. ORCID for DSpace (2014). http://www.slideshare.net/bramluyten/orcid-identifiers-for-dspace

Myntti, J. & Cothran N. (2013). Authority Control in a Digital Repository: Preparing for Linked Data. *Journal of Library Metadata*. Volume 13, Issue 2-3, 2013.

Tarver, H., Waugh, L., Philips, M. & Hicks W. (2013). Implementing name authority control into institutional repositories: A staged approach. UNT Digital Library. http://digital.library.unt.edu/ark:/67531/metadc172365/.

University of Central Florida Libraries (2015). STARS: Showcase of Text, Archives, Research & Scholarship. https://library.ucf.edu/find/stars-digital-repository/

Cover background photo: WorldCat Works linked data

Thank you!

Contact: Sai Deng, Metadata Librarian, sai.deng@ucf.edu

Acknowledgements:

- Lee Dotson, Digital Initiatives Librarian, Lee.Dotson@ucf.edu

- Peter Spyers-Duran, Cataloging Librarian, Peter.Spyers-Duran@ucf.edu

- Jeanne Piascik, Principal Cataloger, Jeanne.Piascik@ucf.edu

- Mary Page, Associate Director, Collections & Technical Services, Mary.Page@ucf.edu