LINKED DATA LEARNING BY DOING

at Michigan State University Libraries

Robin Dean / Digital Projects Librarian
Joshua Barton / Head of Cataloging & Metadata Services
Formed in 2015
Staff from two Divisions:
  • Technical Services
  • Digital Information & Systems
Initial goals:
  • Build expertise
  • Engage in “learn-by-doing”

Two quarter-time reports to Head of Cataloging
Five additional at-will participants
• Striving for right-sized engagement
• No mandate to operationalize
• Instead, proactively assess
• Be informed; be prepared to inform others
• Differentiate between practical potential and hype

TEAM APPROACH

• An approach that is “skeptical”, experimental, and assessment-oriented
Phase 1: Learning

- Monthly meetings to discuss book chapters and exercises
- Good way to learn more about each others’ interest and skills

Linked Data for Libraries, Archives, and Museums
by Seth van Hooland and Ruben Verborgh
http://book.freeyourmetadata.org/
Phase 2: Doing

Project criteria

• Not something we could easily do manually
• Scope was large enough to show potential of linked data
• Scope was small enough to do with current resources
• Allowed us to use a linked-data specific technology
• Relevant to cataloging data
Phase 2: Doing

1. Install triplestore software
2. Load triples into it
3. Query the triples using SPARQL
Project CATNIP

Successes

• Triple store: Apache Jena on a virtual machine
• Loaded entire LOC name authority file (MADS and SKOS)
• Able to query triples in Apache Fuseki (SPARQL server)
Project CATNIP

Lessons Learned

• It’s hard to work with huge text files (10GB +)
• It helps to know what you’re going to use the data for
Project CATNIP

Next Steps

• Name reconciliation (personal or corporate?)
• Find related URIs and context from other vocabularies
• Convert MARC records with future BIBFRAME 2.0 converter
QUESTIONS?

Robin Dean – rdean@msu.edu
Joshua Barton – bartonjp@msu.edu