“Linked-Data-Ready” Software For Libraries: The eXtensible Catalog (XC)

Jennifer Bowen, University of Rochester
CNI Spring Membership Meeting, 2012
April 3, Baltimore, MD
Why should we do it?
Who should do it?
How can we get started?
What are the possible outcomes?

...vis-à-vis the eXtensible Catalog
Linked Data for Libraries: Why?

What do our users need, and can linked data help us provide that?

Are there NEW ROLES that Linked Data can help libraries fulfill?
Studying scholars at the UR...
Scholars want to read everything on the topic that they are researching.
They want to be in the middle of everything they need, all organized so it is findable and usable.
Scholars want their research to be findable and usable by others.
Scholars want to connect to people whose work is interesting and useful to them.
Scholars don’t care *what* the technology is, as long as it helps them do their work.
A shift in how people seek and use information

Systems that libraries provide (websites, catalogs, databases) are bypassed
…not just in favor of Google and the Web in general
…but also in favor of tailored desktop, mobile, and web applications
Beyond library finding tools

“Even scholars who continue to use library finding tools are turning to new applications to aggregate and analyze information in ways that extend their scholarship beyond what manual searching and analyzing allows.”

-- Nancy Fried Foster
   Director of Anthropological Research
   UR River Campus Libraries
Make library resources discoverable on the open web, through applications that potential readers are already using:

- Search engines
- Mobile apps
- Social media
An example...
An example…Mt. Hope Cemetery

Photo credits: ROCHESTER’S SPEAKING STONES By Th. Emil Homerin; University of Rochester Department of Religion and Classics  http://www.rochester.edu/College/REL/faculty/homerin/REL167/reports.htm
An example... Mt. Hope Cemetery

Mt Hope Cemetery, near Rochester, NY

Mount Hope Cemetery
Rochester, New York

Photo credit: www.findagrav.com/cgi-bin/fg.cgi?page=pv&GRid=31&Plpi=76016

Letter of William Channing Gannett describing Susan B. Anthony’s funeral.

(page 2) The services yesterday were simple, heart felt & impressive--college girls acting as "Honorary Bearers" and other girls as Guard of Honor as she lay in State in the Church while the people for three hours streamed through to look upon her face. The congregation filled [Central Church] --one of the largest in the city--to the doors & windows, while others stood outside in falling snow for the hour or two the service lasted...It does not seem like death, does it, --to move people so...
Tools like this are possible today with dedicated programming.

Linked Data will enable library resources to be included in applications like this by allowing application developers access to a “…a store of machine-actionable data on which improved services can be built”. (LOD value statement)
WHO should create library linked data?  

As many libraries as possible.
Libraries need to embrace a new paradigm – a new way of thinking about data.

-- We need *hands-on experience* with linked data to understand linked data’s potential, and to develop best practices for linked data.
Why should a library create linked data?

Serve the unique needs of local users:

- Create linked data for local resources (IR, etc.)
- Showcase special collections
- Develop discovery tools to address needs of local users
Encourage vendors to implement linked data

Educate them about what linked data is, why it is important to libraries.

Get them thinking about/working on linked data NOW, not five years from now!
Why should a library create linked data?

Create/take advantage of new opportunities, new roles for library expertise

LOD value statement: “...opportunities for creative innovation in digital scholarship and participation.”
How can we get started?
To create linked data, we need...

... a tool or tools that will enable us to move from legacy library metadata to library linked data.

Can eXtensible Catalog software play a part in this?
What is XC software?

**eXtensible Catalog (XC)** is open source software for libraries.

XC provides a **discovery system** and a **set of tools** for libraries to manage/transform metadata and build applications.

Four software toolkits available at: [www.eXtensiblecatalog.org](http://www.eXtensiblecatalog.org)
Major Funding
Andrew W. Mellon Foundation

Sponsors
Consortium of Academic and Research Libraries in Illinois (CARLI)
Kyushu University
University of North Carolina at Charlotte
University of Rochester
Welcome to Perseus 4.0, also known as the Perseus Hopper.
Read more on the Perseus version history.
New to Perseus? Click here for a short tutorial.

Announcements

• March 20, 2012

Perseus announces plans to decentralize the curation, annotation, and general editing of texts that it hosts. Ultimately this will include every textual object in Perseus, allowing it to modify (where rights allow), and to create new, dictionary and encyclopedia entries, transcriptions, commentaries, introductions, as well as machine actionable annotations such as identifying people and places and the morpho-syntactic analyses in the Greek and Latin Treebanks. In demos of the prototype interface for such basic tasks as editing and reviewing existing sources or creating new translations. Our goal is to integrate this work with existing learning management systems, evolving e-portfolios, institutional repositories and new instruments of digital pedagogy. We hope to be able to begin beta testing the environment soon.

• January 3, 2012
Linked Data in XC
XC provides a platform for experimentation with metadata transformation/reuse

– and potentially for linked data.
Bulk conversion of existing library metadata
Synchronize data conversion to existing systems
Provides a risk-free way to experiment with data
Can potentially make linked data available to developers in formats that they need
XC software

Drupals Toolkit
- User Interface
  - Search
  - Browse

MST Toolkit
- Metadata Services
  - Cleanup
  - Format Convert

OAI Toolkit
- ILS Connectivity
- Synchronize data with XC

NCIP Toolkit
- ILS Connectivity
  - Circ. status
  - Account info

Digital Repository

Voyager ILS
- Voyager “Driver”

User Interface

Metadata

Live Circ. Data
 XC as Linked-Data-Ready: Options

Linked Data to be created through XC’s bulk metadata conversion processes:
- RDF/XML: harvestable record sets
- SPARQL Endpoint: query for information on demand

Linked Data to be created through Drupal 7 User Interface functionality:
- RDFa
XC: Making Library Metadata
“Linked Data Ready”
The underlying schema for XC uses elements from registered element sets to facilitate conversion to RDF triples (i.e. they already have URIs)
This resource has subject Poets, American.
XC Schema Properties

Dublin Core terms (all)
RDA – subset of elements and role designators
XC elements (newly-defined) – when necessary

All properties are from registered element sets and thus already have URIs
XC converts MARC data to FRBR entities as an interim step. This may enable us to produce more *meaningful* linked data.
“FRBRized” MARC records

Parsing MARCXML records into linked FRBR-based XC Schema records
This resource has subject Poets, American
Without FRBR:

<MARCBibRecord-number> has_author “J K Rowling”

With FRBR:

<Work-id> has_creator “J K Rowling”
<Expression-id> has_language “English”
<Expression-id> has_parent_work <Work-id>
<Manifestation-id> has_isbn <ISBN-number>
<Manifestation-id> has_parent_expression <Expression-id>
Why use FRBR for linked data?

User research shows that users want to see the relationships between resources, etc. With XC, we can explore when/how FRBR might be useful for linked data. Other data models may be more appropriate in some contexts and those can be explored as well.
Possible Outcomes
Revisiting user research findings about scholars

Connect scholars with others whose work is interesting and useful to them

Make scholars’ research findable and usable by others

Develop technology that helps scholars do their work
Create tools that use linked data to...

Allow scholars to create linked data as part of the scholarly process

Follow citation relationships between resources and their authors, etc.

Create and manage vocabularies

Enable experts, etc. to augment metadata about a resource/dataset to make it more discoverable or understandable.
“My thesis is based on this dataset.”
“These photographs are all of the same person”

Photo credits: University of Rochester Frederick Douglass Institute and University of Rochester River Campus Libraries Department of Rare Books and Special Collections. [http://www.lib.rochester.edu/index.cfm?page=2883](http://www.lib.rochester.edu/index.cfm?page=2883)
These other researchers cite MY research...
Open source software offers risk-free experimentation with MARC (and other) library metadata, data manipulation

Potential for bulk creation of linked data in three different ways (RDF/XML, RDFa, SPARQL)

Platform for development of Linked Data tools, to create new opportunities for libraries
Next Steps for XC and Linked Data

We are now seeking funding for more open source software development related to linked data and the needs of scholars.

We invite libraries to participate/partner with us.