Exposing the UIUC Library Catalog as Linked Open Data

Presenters: M. Janina Sarol
Timothy W. Cole

Co-Authors: Myung-Ja Han
Patricia Lampron

University of Illinois at Urbana-Champaign
To share a snapshot of our library catalog including detailed holdings information

• using LOD to maximize usefulness & support emerging services

• providing us an opportunity to assess & quantify challenges of transforming our catalog records into LOD graphs

• and an opportunity to learn about & experiment with the integration & implementation of LOD services & interfaces
MODS with Linked Data: WorldCat

OCLC Number in MARCXML

```xml
<marc:datafield tag="035" ind1=" " ind2=" ">
  <marc:subfield code="a">(OCoLC)ocm00160100</marc:subfield>
</marc:datafield>
```

WorldCat Link

http://worldcat.org/oclc/00160100

MODS

```xml
<identifier type="uri" displayLabel="WorldCat Linked Data">
  http://www.worldcat.org/oclc/00160100
</identifier>
```
• 1.57 names per record
• 1.04 VIAF links per record

Milford, Humphrey Sumner, Sir 1877-1952

<table>
<thead>
<tr>
<th># of Names in Record</th>
<th># of Records</th>
<th># of Records with at least 1 VIAF link</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 names</td>
<td>212,214</td>
<td>0</td>
</tr>
<tr>
<td>1 name</td>
<td>3,393,227</td>
<td>2,291,255</td>
</tr>
<tr>
<td>2 names</td>
<td>1,174,123</td>
<td>1,011,631</td>
</tr>
<tr>
<td>3 or more names</td>
<td>700,185</td>
<td>669,112</td>
</tr>
<tr>
<td>TOTAL</td>
<td>5,479,749</td>
<td>3,971,998</td>
</tr>
</tbody>
</table>
MODS with Linked Data: VIAF

- 3,184,748 unique name strings
- 1,698,218 unique VIAF links

http://viaf.org/24599809/
Coleridge, Samuel Taylor 1772-1834
Coleridge, Samuel Taylor

http://viaf.org/24599809/
Wolcot, John, 1738-1819
Pindar, Peter
LCSH concepts can be simple or complex

- Simple types are a single word or phrase, e.g.: *Vietnam War, 1961-1975*

  <subject authority="lcsh">
    <topic authorityURI="http://id.loc.gov" valueURI="http://id.loc.gov/authorities/subjects/sh85143277"/>
  </subject>

- Complex types concatenate simple, e.g.: *Songs, English -- History and criticism*

  <subject authority="lcsh" authorityURI="http://id.loc.gov"
    valueURI="http://id.loc.gov/authorities/subjects/sh2010114156">
    <topic authorityURI="http://id.loc.gov" valueURI="http://id.loc.gov/authorities/subjects/sh85125063"/>
    <topic authorityURI="http://id.loc.gov" valueURI="http://id.loc.gov/authorities/subjects/sh99001187"/>
  </subject>

URIs exist for some but not all complex type headings

URI of subdivision form used for terms in 2nd, 3rd, ... position of complex types
MODS with Linked Data: LCSH

Subject Types

- Simple Topic: 71%
- Simple Other: 7%
- Complex Subject: 22%

Simple Topics (22%)
- Found: 95%
- Not Found: 5%

Complex Subjects (71%)
- Found: 41%
- Not Found: 59%
**Obvious mappings to Schema.org**

<table>
<thead>
<tr>
<th>MODS</th>
<th>Schema.org</th>
</tr>
</thead>
<tbody>
<tr>
<td>titleInfo</td>
<td>name</td>
</tr>
<tr>
<td>languageTerm</td>
<td>inLanguage</td>
</tr>
<tr>
<td>classification</td>
<td>about</td>
</tr>
<tr>
<td>genre</td>
<td>genre</td>
</tr>
<tr>
<td>dateModified</td>
<td>dateModified</td>
</tr>
<tr>
<td>dateCreated</td>
<td>dateCreated</td>
</tr>
<tr>
<td>copyrightDate</td>
<td>copyrightYear</td>
</tr>
<tr>
<td>publisher</td>
<td>publisher</td>
</tr>
<tr>
<td>place</td>
<td>placeOfPublication (library extension)</td>
</tr>
</tbody>
</table>
### Mapping names

<table>
<thead>
<tr>
<th>MODS</th>
<th>Schema.org</th>
</tr>
</thead>
<tbody>
<tr>
<td>name</td>
<td>Person Organization</td>
</tr>
<tr>
<td>role</td>
<td>author</td>
</tr>
<tr>
<td></td>
<td>contributor</td>
</tr>
<tr>
<td></td>
<td>creator</td>
</tr>
<tr>
<td></td>
<td>editor</td>
</tr>
<tr>
<td></td>
<td>illustrator</td>
</tr>
</tbody>
</table>

```xml
<sc:creator>
  <sc:Person rdf:about="http://viaf.org/viaf/76492614/"/>
</sc:creator>.
<sc:creator>
  <sc:Organization>
    <sc:name>Oxford University Press</sc:name>
  </sc:Organization>
</sc:creator>
<sc:contributor>
  <sc:Person>
    <sc:name>Gray, Austin K.</sc:name>
    <sc:jobTitle>donor</sc:jobTitle>
  </sc:Person>
</sc:contributor>
```
# Mapping LCSH subjects

<table>
<thead>
<tr>
<th>MODS</th>
<th>Schema.org</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject</td>
<td>about</td>
</tr>
</tbody>
</table>

## Subject Types

<table>
<thead>
<tr>
<th>MODS</th>
<th>madsrdf</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic</td>
<td>Topic</td>
</tr>
<tr>
<td>Geographic</td>
<td>Geographic</td>
</tr>
<tr>
<td>Name</td>
<td>PersonalName</td>
</tr>
<tr>
<td>TitleInfo</td>
<td>Title</td>
</tr>
<tr>
<td>Temporal</td>
<td>Temporal</td>
</tr>
<tr>
<td>Genre</td>
<td>GenreForm</td>
</tr>
</tbody>
</table>

## Statistics

```xml
<sc:about>
  <madsrdf:Topic rdf:about="http://id.loc.gov/authorities/subjects/sh85127580"/>
</sc:about>
```

## France--Description and Travel

```xml
<sc:about>
  <madsrdf:ComplexSubject rdf:about="http://id.loc.gov/authorities/subjects/sh85051207">
    <sc:hasPart>
      <madsrdf:Geographic>
        <sc:name>France</sc:name>
      </madsrdf:Geographic>
    </sc:hasPart>
    <sc:hasPart>
      <madsrdf:Topic rdf:about="http://id.loc.gov/authorities/subjects/sh99001278"/>
    </sc:hasPart>
  </madsrdf:ComplexSubject>
</sc:about>
```
Use Case: HTML with RDFa
Use Case: HTML with RDFa
Observations & future plans

Observations:

• Strings into URIs hard. Issues: scale, completeness, data quality & ambiguity. More / better reconciliation services are needed.
• There are competing semantics for mapping to RDF. Community needs to coalesce.
• Models & services that leverage LOD in libraries are immature. Need more exemplars & experimentation.

Looking ahead for UIUC:
• Finish releasing RDF for all records in our catalog.
• Continue to explore & experiment with new services & interface designs
http://catalogdata.library.illinois.edu/

Janina Sarol (mjsarol@illinois.edu)
Tim Cole (t-cole3@illinois.edu)
Myung-Ja Han (mhan3@illinois.edu)
Patricia Lampron (lampron2@illinois.edu)