What do MARC, RDF, and OWL have in common?

Violeta Ilik
Head, Digital Systems & Collection Services
Galter Health Sciences Library

Program for Cooperative Cataloging (PCC) Participants Meeting - ALA 2015, San Francisco, June 28, 2015
Catalogers and MARC

Feels like home!

• You know the rules,
• You know all the intricacies
Catalogers and XML, XSLT, namespaces, MarcEdit ...

Feels like home?
• Do you know the rules?
• Do you know all the intricacies?

Crosswalk:

Mapping compares and analyzes two or more metadata schemas, while crosswalks are the product of the mapping process.

Each XSLT stylesheet describes how a set of XML documents (the source documents) should be converted to other documents (the result documents).
MARCXML, Dublin Core, XML, BIBFRAME (JSON)
Catalogers and repository metadata

Should feel like home!

- Descriptive, administrative and structural metadata
- Use of established controlled vocabularies
- Use of institutional databases – campus directory
- Use of established standards (date, language)
"ontology is an explicit representation of conceptualization"
Gruber, 1993

“RDF is a standard model for data interchange on the Web.” w3c.org/RDF/
Catalogers and repository metadata

Should feel like home!

• Descriptive, administrative and structural metadata
• Use of established controlled vocabularies
• Use of university databases – campus directory
• Use of established standards (date, language)
• Using ontologies

Existing options for resource type in Fedora/Sufia
Protégé - understanding ontologies

According to Crofts, Le Boef, and Artur (2002):

- An ontology is a *formal* representation which aims to be precise, explicit and unambiguous.
- The ontology describes the entities and concepts relevant to a particular *domain*.
- An ontology embodies a particular *view* of a domain - the same domain may be described in different ways by different ontologies.
Catalogers and repository metadata

Should feel like home!

• Descriptive, administrative and structural metadata
• Use of established controlled vocabularies
• Use of university databases – campus directory
• Use of established standards (date, language)
• Using ontologies
• Creating RDF data

From VIVO

Fedora/Sufia Repository record
RDF for VIVO: from CSV to RDF via Karma

<table>
<thead>
<tr>
<th>UID</th>
<th>NETID</th>
<th>Full Name</th>
<th>First Name</th>
<th>Last Name</th>
<th>Mid Name</th>
<th>Email</th>
<th>date_hired</th>
<th>date_terminated</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2dec06e9</td>
<td>Sadav</td>
<td>Sadav, Manjit S.</td>
<td>Manjit</td>
<td>Sadav</td>
<td>s</td>
<td><a href="mailto:Sadav@school.ed">Sadav@school.ed</a></td>
<td>2000-10-01T00:00:00</td>
<td>2014-01-01T00:00:00</td>
<td>Professor</td>
</tr>
<tr>
<td>c81ef444</td>
<td>mindyGergman</td>
<td>Gergman, Mindy E.</td>
<td>Mindy</td>
<td>Gergman</td>
<td>e</td>
<td>mindyGergman@sch...</td>
<td>2000-10-02T00:00:00</td>
<td>2014-01-02T00:00:00</td>
<td>Associate Professor</td>
</tr>
<tr>
<td>5c2ce0e</td>
<td>t_Vay</td>
<td>Vay, Matilda W.</td>
<td>Matilda</td>
<td>Vay</td>
<td>w</td>
<td><a href="mailto:t_Vay@school.ed">t_Vay@school.ed</a></td>
<td>2000-10-03T00:00:00</td>
<td>2014-01-03T00:00:00</td>
<td>Lecturer</td>
</tr>
<tr>
<td>b6f25071</td>
<td>sBach</td>
<td>Bach, Susan A.</td>
<td>Susan</td>
<td>Bach</td>
<td>a</td>
<td><a href="mailto:sBach@school.ed">sBach@school.ed</a></td>
<td>2000-10-04T00:00:00</td>
<td>2014-01-04T00:00:00</td>
<td>Clinical Associate Professor</td>
</tr>
<tr>
<td>613870d1</td>
<td>kRoberts</td>
<td>Roberts, Kathryn J.</td>
<td>Kathryn</td>
<td>Roberts</td>
<td>j</td>
<td>kRoberts@sch...</td>
<td>2000-10-05T00:00:00</td>
<td>2014-01-05T00:00:00</td>
<td>Assistant Professor</td>
</tr>
<tr>
<td>980e771d</td>
<td>jasenMaldonado</td>
<td>Maldonado, Jasen J.</td>
<td>Jasen</td>
<td>Maldonado</td>
<td>j</td>
<td>jasenMaldonado@sch...</td>
<td>2000-10-06T00:00:00</td>
<td>2014-01-06T00:00:00</td>
<td>Assistant Professor</td>
</tr>
<tr>
<td>ea52a51f</td>
<td>Sahrrana20</td>
<td>Sahrrana, Krishan C.</td>
<td>Krishan</td>
<td>Sahrrana</td>
<td>c</td>
<td><a href="mailto:Sahrrana20@school.ed">Sahrrana20@school.ed</a></td>
<td>2000-10-07T00:00:00</td>
<td>2014-01-07T00:00:00</td>
<td>Visiting Associate Professor</td>
</tr>
</tbody>
</table>

---

Violeta Ilik
ALA PCC Participants Meeting, June 28, 2015
San Francisco
N-Triples — VIVO compliant data
We want to optimize discoverability and dissemination of content and enhance the impact of FSM, NUCATS, and our Northwestern Medicine community.
Other projects led by Digital Systems & Collection Services

Symplectic Elements

- Back-end bibliometric aggregator
- Support OA with repository integration
- Facilitates reports and reuse of clean aggregated data from a number of diverse sources
Symplectic elements

<table>
<thead>
<tr>
<th>Metadata</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Altmetric</td>
<td>-</td>
</tr>
<tr>
<td>arXiv</td>
<td>✓</td>
</tr>
<tr>
<td>CiNii</td>
<td>✓</td>
</tr>
<tr>
<td>CrossRef*</td>
<td>✓</td>
</tr>
<tr>
<td>dblp</td>
<td>✓</td>
</tr>
<tr>
<td>Europe-PMC</td>
<td>✓</td>
</tr>
<tr>
<td>figshare</td>
<td>✓</td>
</tr>
<tr>
<td>Google Books</td>
<td>✓</td>
</tr>
<tr>
<td>PubMed</td>
<td>✓</td>
</tr>
<tr>
<td>RePEc</td>
<td>✓</td>
</tr>
<tr>
<td>Scopus*</td>
<td>✓</td>
</tr>
<tr>
<td>Web of Science*</td>
<td>✓</td>
</tr>
</tbody>
</table>

Modeling Symplectic elements’ publications data

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>uuid</td>
<td>ID</td>
<td>Title</td>
<td>Journal ORP year</td>
<td>DOI</td>
<td>Volume</td>
<td>Issue</td>
<td>ISSN</td>
<td>pages</td>
<td>Pagination (start page)</td>
<td>Pagination (end page)</td>
<td>Publication Date</td>
<td>Person Name</td>
<td>abstract</td>
</tr>
<tr>
<td>cwid-jon9024</td>
<td>pubid25043</td>
<td>Does adjuvant Cancer</td>
<td>2014</td>
<td>10.1002/cnci</td>
<td>120</td>
<td>22 0008-543X</td>
<td>6</td>
<td>3562</td>
<td>3568</td>
<td>2014-11-15T00:00:00</td>
<td>J. Ng</td>
<td>A characteristic feature of the b</td>
<td></td>
</tr>
</tbody>
</table>

[Diagram and table content]
Projects to which we contribute:
VIVO @ Northwestern FSM
Projects to which we contribute: VIVO @ Northwestern FSM

Data modeling using the Karma data integration tool

R2RML (W3C standard) models that can be re-applied to same data source (CSV export files from data stores)

Ontology development
What do MARC, OWL, and RDF have in common?

• They all need a cataloger, or in other words – one honest trouble maker
Thank you

Violeta Ilik

vivoweb.org