# 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

NORTHWESTERN UNIVERSITY
**FEINBERG**
SCHOOL OF MEDICINE

# Catalogers and MARC

## Feels like home!

- You know the rules,

- You know all the intricacies

```
LEADER 00000ngm  2200349Ka 4500
001    ocn271676443
003    OCoLC
007    vf czahzu
008    081114s2008    xx 059          vleng d
040    DRB|cDRB
049    DRBB
245 04 The future of humanities scholarship in a digital world
       |h[videorecording].
260    |c2008.
300    1 videocassette (59 min.) :|bsd., col. ;|c1/5 in.
500    Title from container.
500    Viewing copies available on DVD.
511 0  Laura Mandell.
518    Recorded on August 1, 2008, in 105 Dartmouth Hall,
       Dartmouth College.
538    DVCam.
655  7 College films and programs.|2mim
700 1  Mandell, Laura.
710 2  Dartmouth College.|bMedia Production Group.
710 2  Dartmouth College.|bResearch Computing.
948    LTI 12/09/2008
955    |u+l
998    |cjat/lgg
```

**MARC record**

# 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)

Violeta Ilik
ALA PCC Participants Meeting, June 28, 2015
San Francisco

# MARCXML, Dublin Core XML, BIBFRAME (JSON)

```
<datafield tag="245" ind1="0" ind2="4">
  <subfield code="a">The future of humanities scholarship in a digital world</subfield>
  <subfield code="h">[videorecording].</subfield>
</datafield>
<datafield tag="260" ind1=" " ind2=" ">
  <subfield code="c">2008.</subfield>
</datafield>
<datafield tag="300" ind1=" " ind2=" ">
  <subfield code="a">1 videocassette (59 min.) :</subfield>
  <subfield code="b">sd., col. ;</subfield>
  <subfield code="c">1/5 in.</subfield>
</datafield>
<datafield tag="500" ind1=" " ind2=" ">
  <subfield code="a">Title from container.</subfield>
</datafield>
<datafield tag="511" ind1="0" ind2=" ">
  <subfield code="a">Laura Mandell.</subfield>
</datafield>
<datafield tag="518" ind1=" " ind2=" ">
  <subfield code="a">Recorded on August 1, 2008, in 105 Dartmouth Hall, Dartmouth
  College.</subfield>
</datafield>
<datafield tag="500" ind1=" " ind2=" ">
  <subfield code="a">Viewing copies available on DVD.</subfield>
</datafield>
<datafield tag="538" ind1=" " ind2=" ">
  <subfield code="a">DVCam.</subfield>
</datafield>
<datafield tag="655" ind1=" " ind2="7">
  <subfield code="a">College films and programs.</subfield>
  <subfield code="2">mim</subfield>
</datafield>
<datafield tag="700" ind1="1" ind2=" ">
  <subfield code="a">Mandell, Laura.</subfield>
</datafield>
<datafield tag="710" ind1="2" ind2=" ">
  <subfield code="a">Dartmouth College.</subfield>
  <subfield code="b">Media Production Group.</subfield>
</datafield>
<datafield tag="710" ind1="2" ind2=" ">
  <subfield code="a">Dartmouth College.</subfield>
  <subfield code="b">Research Computing.</subfield>
</datafield>
</record>
```

```
<rdf:Description>
    <dc:title>The future of humanities scholarship in a digital world [videorecording]</dc:title>
    <dc:creator>Mandell, Laura</dc:creator>
    <dc:contributor>Dartmouth College. Media Production Group</dc:contributor>
    <dc:contributor>Dartmouth College. Research Computing</dc:contributor>
    <dc:type>College films and programs. mim</dc:type>
    <dc:publisher/>
    <dc:date.created>2008</dc:date.created>
    <dc.description>1 videocassette (59 min.) : sd., col. ; 1/5 in.</dc.description>
    <dc.description>Title from container.</dc.description>
    <dc.description>Viewing copies available on DVD.</dc.description>
    <dc:type>moving image</dc:type>
    <dc:language.iso>en_US</dc:language.iso>
    <dc:type.material>video</dc:type.material>
    <dc:format.digitalOrigin>reformatted digital</dc:format.digitalOrigin>
    <dc:format.medium>electronic</dc:format.medium>
  </rdf:Description>
</rdf:RDF>
```

```
type: "Instance",
id: "ocn271676443instance14",
bf-id: "ocn271676443instance14",
uri: "http://bibframe.org/resources/IIn1435410184/ocn271676443instance14",
label: "future of humanities scholarship in a digital world",
bf-instanceTitle: "ocn271676443title17",
bf-title: "future of humanities scholarship in a digital world",
bf-publication: [
  - {
        type: "Provider",
        bf-copyrightDate: "2008"
    }
],
bf-modeOfIssuance: "single unit",
bf-extent: "1 videocassette (59 min.) :",
bf-dimensions: "1/5 in.",
bf-illustrationNote: "sd., col. ;",
bf-titleStatement: "The future of humanities scholarship in a digital world",
bf-formDesignation: "videorecording.",
bf-providerStatement: "2008.",
bf-performerNote: "Laura Mandell.",
bf-note: [
    "Title from container.",
    "Viewing copies available on DVD.",
    "DVCam."
],
bf-derivedFrom: "ocn271676443.marcxml.xml",
bf-instanceOf: "ocn271676443"
```

# 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)



Fedora/Sufia Repository record

Violeta Ilik
ALA PCC Participants Meeting, June 28, 2015
San Francisco

# OWL and RDF

**"ontology is an explicit representation of conceptualization"**
Gruber, 1993

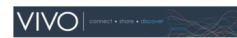**"RDF is a standard model for data interchange on the Web." w3c.org/RDF/**

# 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**

Edit Dos and Don'ts of extending the VIVO ontology

VIVO  connect • share • discover

Dos and Don'ts of extending the VIVO ontology

Violeta Ilik @violetailik
Galter Health Sciences Library
Feinberg School of Medicine
Northwestern University Clinical and
Translational Science Institute (NUCATS), Chicago, IL
https://galter.northwestern.edu/staff/Violeta-Ilik
http://vivo.vivoweb.org/display/n10603

Download the full-sized PDF

| Descriptions | Versions | Permissions |

Descriptions · indicates required fields

**Resource type** ❓
Other
Part of Book
Patient Education Handout
Poster
Presentation

Abstract ❓ [+ Add]

Bibliographic citation ❓ [+ Add]

* Keyword ❓
OWL [– Remove]
VIVO [– Remove]
National Library of Medicine [– Remove]
[+ Add]

* Rights
Attribution-ShareAlike 3.0 United States
Attribution-NonCommercial 3.0 United States
Attribution-NoDerivs 3.0 United States
Attribution-NonCommercial-NoDerivs 3.0 United States
Attribution-NonCommercial-ShareAlike 3.0 United States

```
"Article" => "http://schema.org/Article",
"Audio" => "http://schema.org/AudioObject",
"Book" => "http://schema.org/Book",
"Capstone Project" => "http://schema.org/CreativeWork",
"Conference Proceeding" => "http://schema.org/ScholarlyArticle",
"Dataset" => "http://schema.org/Dataset",
"Dissertation" => "http://schema.org/ScholarlyArticle",
"Image" => "http://schema.org/ImageObject",
"Journal" => "http://schema.org/CreativeWork",
"Map or Cartographic Material" => "http://schema.org/Map",
"Masters Thesis" => "http://schema.org/ScholarlyArticle",
"Part of Book" => "http://schema.org/Book",
"Poster" => "http://schema.org/CreativeWork",
"Presentation" => "http://schema.org/CreativeWork",
"Project" => "http://schema.org/CreativeWork",
"Report" => "http://schema.org/CreativeWork",
"Research Paper" => "http://schema.org/ScholarlyArticle",
"Software or Program Code" => "http://schema.org/Code",
"Video" => "http://schema.org/VideoObject",
"Other" => "http://schema.org/CreativeWork",
```
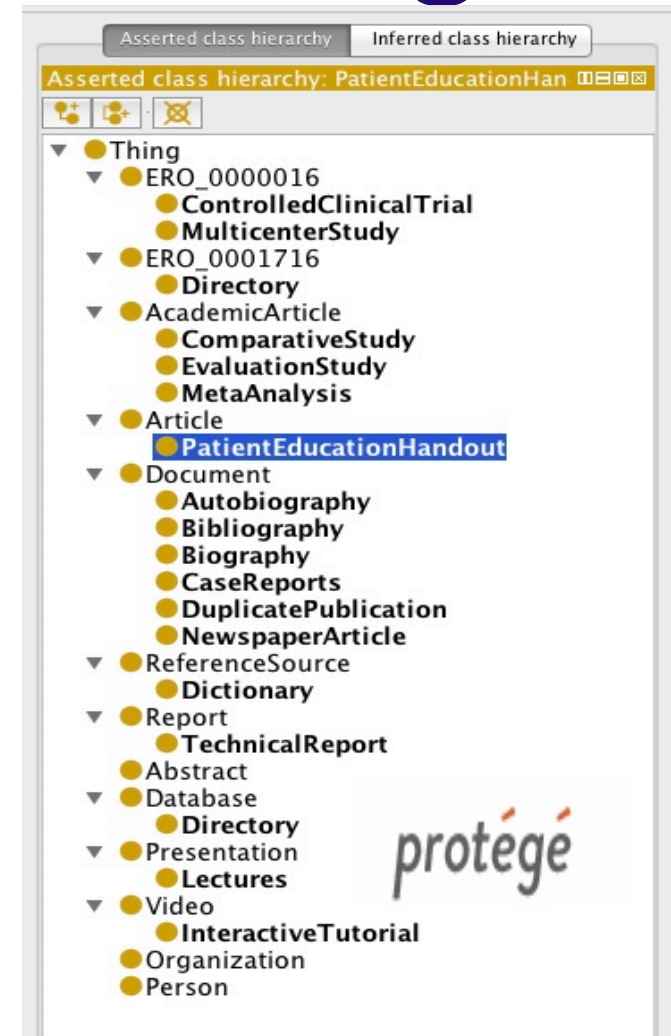
**Existing options for resource type in Fedora/Sufia**

# 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**

Ilik, Violeta | Person

**Fedora/Sufia Repository record**

Violeta Ilik
ALA PCC Participants Meeting, June 28, 2015
San Francisco

# RDF for VIVO: from CSV to RDF via Karma

| UID | NETID | FullName | FirstName | LastName | MidName | Email | date_hired | date_terminated | Title |
|---|---|---|---|---|---|---|---|---|---|
| 2dec06e9 | Sadav | Sadav, Manjit S. | Manjit | Sadav | S | Sadav@scho | 2000-10-01T00:00:00 | 2014-01-01T00:00:00 | Professor |
| c81ef444 | mindyGergman | Gergman, Mindy E. | Mindy | Gergman | E | mindyGergm | 2000-10-02T00:00:00 | 2014-01-02T00:00:00 | Associate Professor |
| 5c2cef0e | t_Vay | Vay, Matilda W. | Matilda | Vay | W | t_Vay@scho | 2000-10-03T00:00:00 | 2014-01-03T00:00:00 | Lecturer |
| b6f25071 | sBach | Bach, Susan A. | Susan | Bach | A | sBach@scho | 2000-10-04T00:00:00 | 2014-01-04T00:00:00 | Clinical Associate Professor |
| 613870d1 | kRoberts | Roberts, Kathryn J. | Kathryn | Roberts | J | kRoberts@sc | 2000-10-05T00:00:00 | 2014-01-05T00:00:00 | Assistant Professor |
| 980e771d | jasenMaldonado | Maldonado, Jasen J. | Jasen | Maldonado | J | jasenMaldon | 2000-10-06T00:00:00 | 2014-01-06T00:00:00 | Assistant Professor |
| ea52a51f | Sahrana20 | Sahrana, Krishan C. | Krishan | Sahrana | C | Sahrana20@ | 2000-10-07T00:00:00 | 2014-01-07T00:00:00 | Visiting Associate Professor |
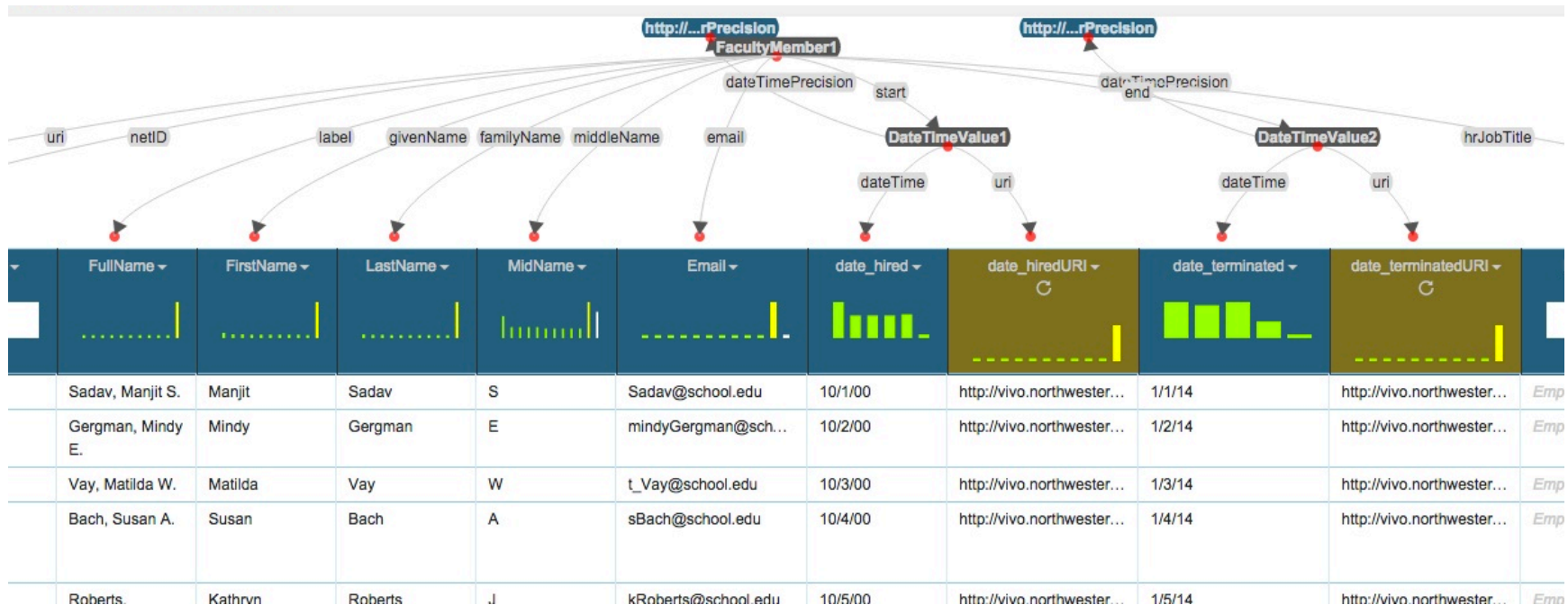
# N-Triples – VIVO compliant data

```
<http://vivo.northwestern.edu/individual/2014-01-01T00:00:00> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://vivoweb.org/ontology/core#DateTimeValue> .
<http://vivo.northwestern.edu/individual/n2dec06e9> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://vivoweb.org/ontology/core#FacultyMember> .
<http://vivo.northwestern.edu/individual/n2dec06e9> <http://www.w3.org/2000/01/rdf-schema#label> "Sadav, Manjit S." .
<http://vivo.northwestern.edu/individual/n2dec06e9> <http://www.w3.org/2006/vcard/ns#givenName> "Manjit" .
<http://vivo.northwestern.edu/individual/n2dec06e9> <http://vivoweb.org/ontology/core#end> <http://vivo.northwestern.edu/individual/2014-01-01T00:00:00> .
<http://vivo.northwestern.edu/individual/n2dec06e9> <http://www.w3.org/2006/vcard/ns#familyName> "Sadav" .
<http://vivo.northwestern.edu/individual/n2dec06e9> <http://vivoweb.org/ontology/core#start> <http://vivo.northwestern.edu/individual/2000-10-01T00:00:00> .
<http://vivo.northwestern.edu/individual/2000-10-01T00:00:00> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://vivoweb.org/ontology/core#DateTimeValue> .
<http://vivo.northwestern.edu/individual/2000-10-01T00:00:00> <http://vivoweb.org/ontology/core#dateTime> "2000-10-01T00:00:00" .
<http://vivo.northwestern.edu/individual/2000-10-01T00:00:00> <http://vivoweb.org/ontology/core#dateTimePrecision> <http://vivoweb.org/ontology/core#yearPrecision> .
<http://vivo.northwestern.edu/individual/2014-01-01T00:00:00> <http://vivoweb.org/ontology/core#dateTimePrecision> <http://vivoweb.org/ontology/core#yearPrecision> .
<http://vivo.northwestern.edu/individual/n2dec06e9> <http://www.w3.org/2006/vcard/ns#email> "Sadav@school.edu" .
<http://vivo.northwestern.edu/individual/n2dec06e9> <http://vivoweb.org/ontology/core#hrJobTitle> "Professor" .
<http://vivo.northwestern.edu/individual/n2dec06e9> <http://vivoweb.org/ontology/core#middleName> "S" .
<http://vivo.northwestern.edu/individual/n2dec06e9> <http://vivo.northwestern.edu/ontology/vlocal#netID> "Sadav" .
<http://vivo.northwestern.edu/individual/2014-01-01T00:00:00> <http://vivoweb.org/ontology/core#dateTime> "2014-01-01T00:00:00" .
```
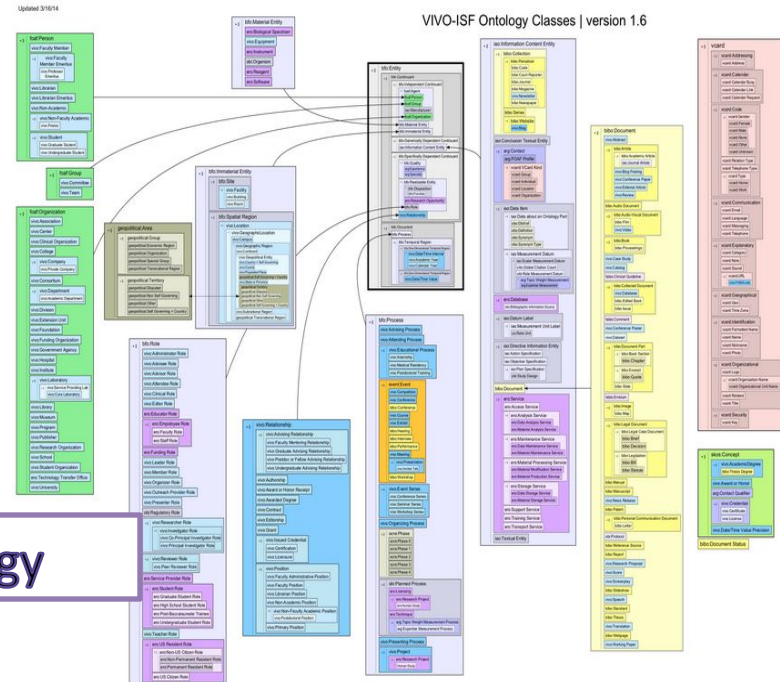
VIVO-ISF Ontology

# Library ecosystem



*We want to optimize discoverability and dissemination of content and enhance the impact of FSM, NUCATS, and our Northwestern Medicine community.*

Violeta Ilik
ALA PCC Participants Meeting, June 28, 2015
San Francisco
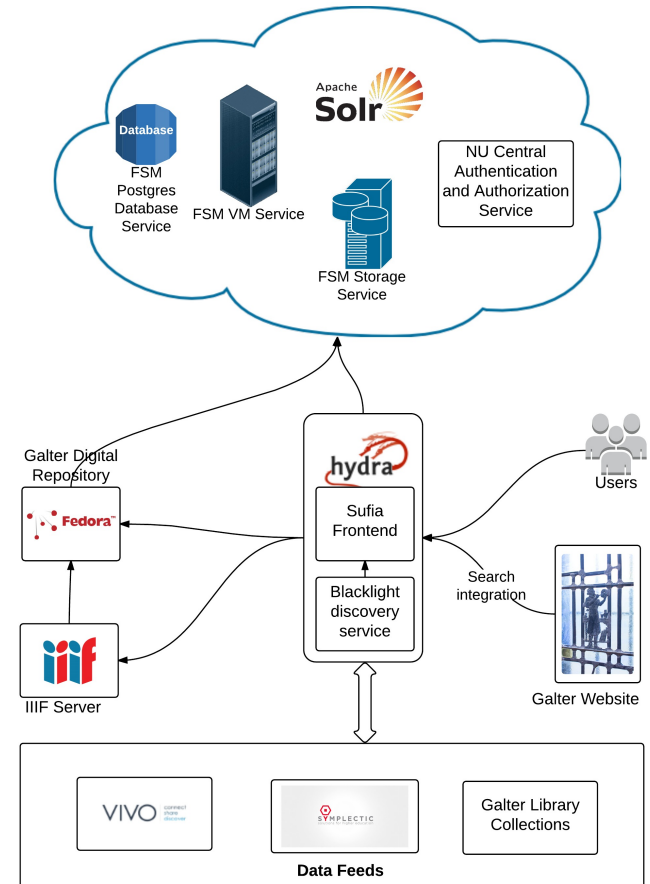
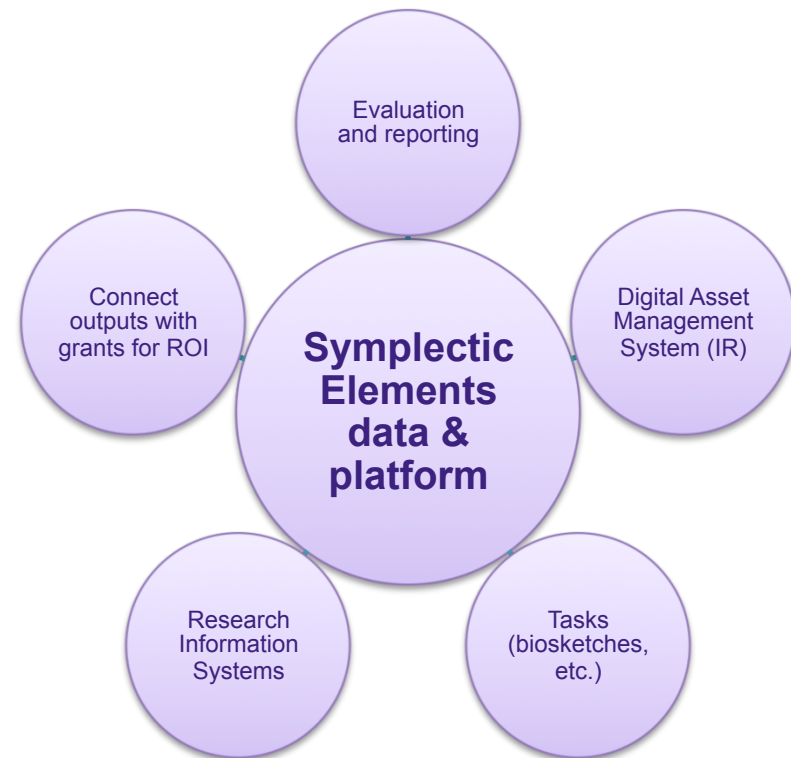# Other projects led by Digital Systems & Collection Services

## 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

| | Metadata |
|---|:---:|
| Altmetric | – |
| arXiv | ✔ |
| CiNii | ✔ |
| CrossRef* | ✔ |
| dblp | ✔ |
| Europe-PMC | ✔ |
| figshare | ✔ |
| Google Books | ✔ |
| PubMed | ✔ |
| RePEc | ✔ |
| Scopus* | ✔ |
| Web of Science* | ✔ |

**Symplectic Elements data & platform**

- Evaluation and reporting
- Digital Asset Management System (IR)
- Tasks (biosketches, etc.)
- Research Information Systems
- Connect outputs with grants for ROI

Holmes, K. 2015. Linked Data Love: research representation, discovery, and assessment. ALCTS/LITA Linked Library Data Interest Group. American Library Association Annual Meeting, San Francisco June 27, 2015

# Modeling Symplectic elements' publications data

# Projects to which we contribute: VIVO @ Northwestern FSM
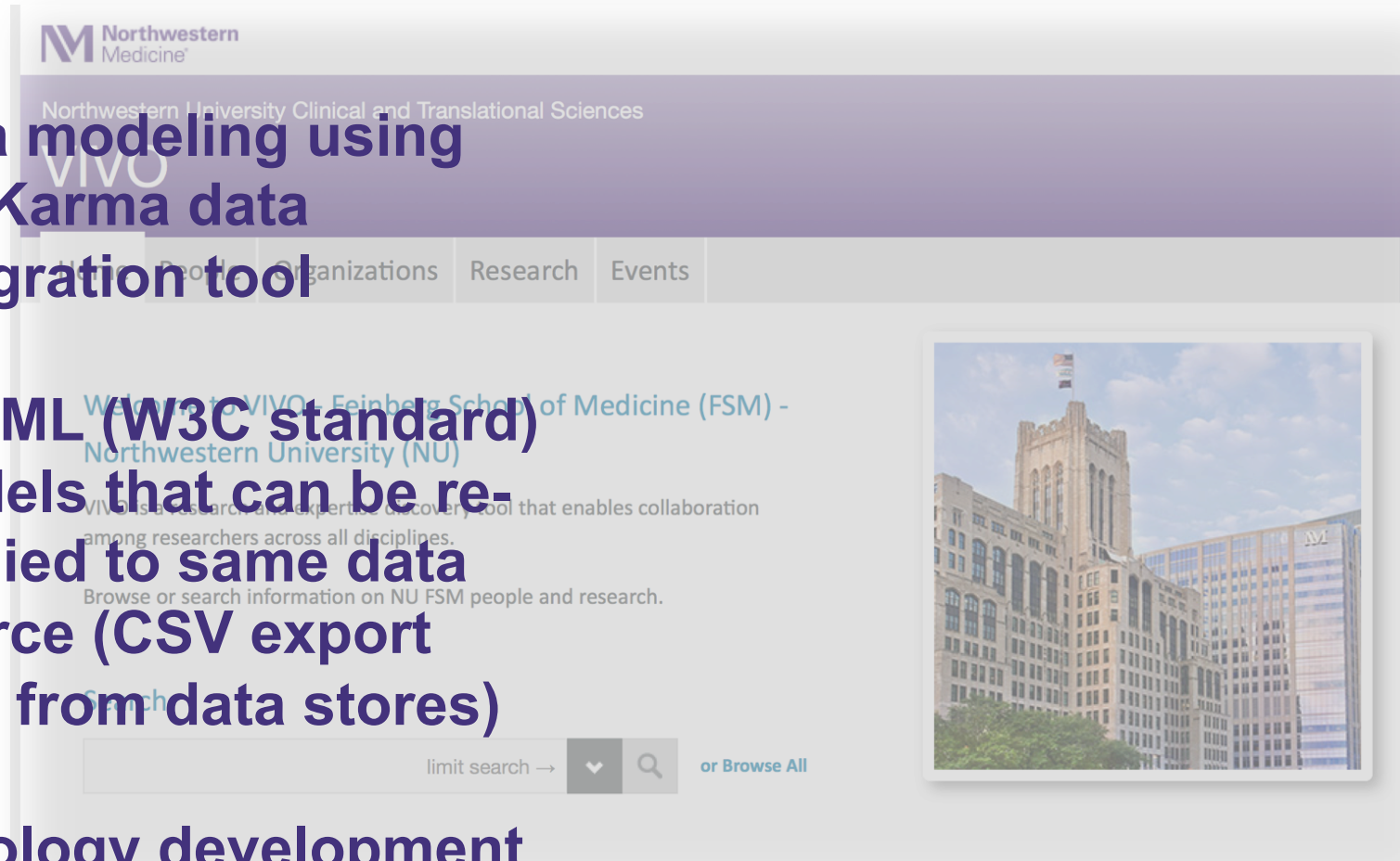
# Projects to which we contribute: VIVO @ Northwestern FSM

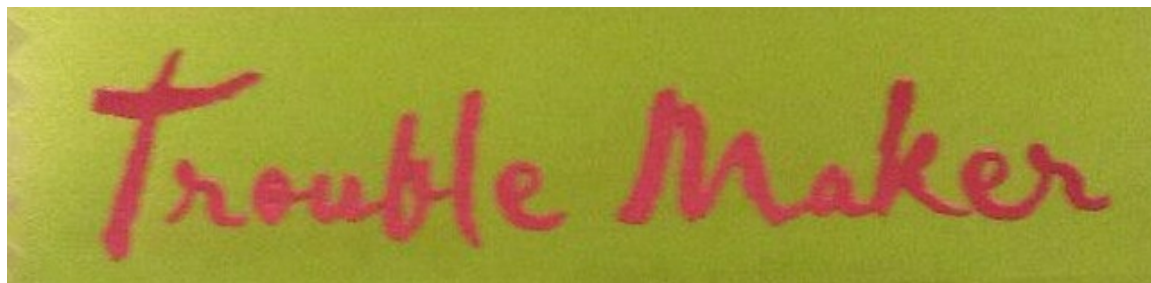**Data modeling using the Karma data integration tool**

**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](vivoweb.org)

iD http://orcid.org/0000-0003-2588-3084

NORTHWESTERN UNIVERSITY
**FEINBERG**
SCHOOL OF MEDICINE