VIVO and the role of librarians
part 2

Violeta Ilik

https://vioil.github.io/
http://openvivo.org/a/orcid0000-0003-2588-3084

Hosted by ALCTS, Association for Library Collections and Technical Services.
About me

https://vioil.github.io/

- Head of Cataloging & Metadata Services at Stony Brook University Libraries, July 2017 - present
- Head, Digital Systems and Collection Services - Northwestern University, Feinberg School of Medicine, Galter Health Sciences Library, Northwestern University Clinical and Translational Sciences Institute, November 2014 – July 2017
- Assistant Professor, Texas A&M University Libraries, February 2012 – October 2014
- Library Specialist I-III, Texas A&M University Libraries, November 2007 - February 2012

Hosted by ALCTS, Association for Library Collections and Technical Services.
Goals for today - part II

- Provide a valuable guide to best practices in modeling RDF data by utilizing data integration tools.
  - Ontology, SPARQL, and all that RDF jazz

Goals for part I

- Gain a deeper understanding of the VIVO semantic web platform.
- Learn how to find the right resources for a new VIVO implementation, including data sources, team members, governance models, and support structures.

Hosted by ALCTS, Association for Library Collections and Technical Services.
Integrating institutional data

Data hub

Organizations

Scholars

Grants/Projects

Publications & Scholarly works

Teaching, Engagement

Websites

Expert Finding

Portfolio/ Vitae

Network Analysis

Reports

Ad hoc Queries

VIVO
<table>
<thead>
<tr>
<th>Provost’s Office</th>
<th>Library</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Resources</td>
<td>Faculty</td>
</tr>
<tr>
<td>Registrar</td>
<td>IT / Information Security</td>
</tr>
<tr>
<td>Office of Research</td>
<td>Other</td>
</tr>
</tbody>
</table>
It's all about data

Things VIVO Talks About

The activities and accomplishments of your people.

Your organization's, people and locations to be included in your VIVO.

Entities to be referred to in your VIVO.
How data gets into VIVO?

1. The user puts it in manually
2. You put it in manually
3. Software puts it in automatically

Hosted by ALCTS, Association for Library Collections and Technical Services.
VIVO is ontologies

- VIVO is ontologies, software, data, and a community for representing scholars and their works
- Ontologies — 1 core ontology (VIVO-ISF), 3 additional VIVO-specific ontologies, 22 related ontologies, including OBO (Open Biological and Biomedical Ontology), BFO (Basic Formal Ontology), and CRO (contribution ontology).

Data in VIVO is stored using Resource Description Framework

Hosted by ALCTS, Association for Library Collections and Technical Services.
Reminder: What is an ontology?

A set of representational primitives with which to model a domain of knowledge or discourse (Tom Gruber, 2007)
VIVO is a knowledge management system that uses RDF standards.

VIVO produces linked data.

The VIVO-ISF ontology provides data properties and object properties used in VIVO assertions.

VIVO-ISF collects data properties and object properties from other ontologies to provide a rich vocabulary for expressing assertions regarding scholarship.

When systems use the VIVO-ISF ontology, they “speak the same language” and data sharing becomes much simpler.

Hosted by ALCTS, Association for Library Collections and Technical Services.
VIVO Classes
Domains of representation
Ontologies used in VIVO

VIVO – [http://vivoweb.org/ontology/core](http://vivoweb.org/ontology/core)
Cell Ontology (CL) – [http://cellontology.org/?q=download](http://cellontology.org/?q=download)
SKOS (Simple Knowledge Organization System) – [http://www.w3.org/2004/02/skos/](http://www.w3.org/2004/02/skos/)
vCard – [http://www.w3.org/TR/vcard-rdf/](http://www.w3.org/TR/vcard-rdf/)
VIVO-ISF Ontology version 1.6

IRI:  
http://vivoweb.org/ontology/core

Other visualisation:
Ontology source

Table of Content

1. Classes
2. Object Properties
3. Data Properties
4. Named Individuals
5. Annotation Properties
6. Namespace Declarations

http://www.essepuntato.it/lode/owlapi
http://vivoweb.org/ontology/core

Hosted by ALCTS, Association for Library Collections and Technical Services.
triples

- VIVO web application is built using "triples"
  1. subject (an entity)
  2. predicate (an object property or a data property)
  3. object (any individual in VIVO)
- These "triples" are also called statements and reflect the structure of a sentence in ordinary language
An Example Regarding Organizations

"College of Science"

"Department of Chemistry"
Statements about the Organizations

Thing1 name “College of Science”.

Thing1 is a College.

Thing1 has part Thing 2.

Thing2 is a Department.

Thing2 name “Department of College”.

These statements are also called assertions. They are also called triples.

See Managing Organization Hierarchy in the VIVO Documentation Wiki
https://wiki.duraspace.org/display/VIVODOC19x/Managing+Organization+Hierarchy
Ontological models

- Organization Model
- Concept Model
- DateTimeValue and DateTimeInterval Models
- Journal Model
- Person Model
- Teaching Model
- Publication Model
- Grant Model
- Education and Training Model
- Advising Model
- Award Model
- Membership Model
- Ontology Diagram Legend
- Credential Model

https://wiki.dura-space.org/display/VIVODOC19x/Ontology+Diagrams

Hosted by ALCTS, Association for Library Collections and Technical Services.
Person

- Has type(s) — faculty, staff, graduate student, etc (these are more correctly modeled as roles)
- Has a vcard for contact information
- Has various identifiers — SCOPUS, ResearcherID, ORCiD, eRACommons
- Photo, home department, research areas (concepts)
- Is related to positions, education, and roles in grants, papers, courses, organizations, and other activities
- The ontologies are sufficient to represent the relationships between the entities and properties of scholars and their works

Hosted by ALCTS, Association for Library Collections and Technical Services.
Journal Model
12 October 2016

Journal
Grant

Note that other people may have other roles on the grant using the same properties but different role classes such as vivo:CoPrincipalInvestigatorRole and vivo:InvestigatorRole.
Education and Training Model
12 October 2016

Education and Training
Advising
Award

Award Model
12 February 2017

Diagram of Award Model with connections and labels.
Membership Model
12 October 2016
DateTimeValue Model

11 October 2016

Date Time
Value
Date Time Interval

DateTimeInterval Model
11 October 2016
Issues

- VIVO occupies the “middle ground” in ontologies
  - more abstract is the BFO and OBO
  - less abstract are domain specific ontologies and vocabularies (clinical trials, agriculture, cultural heritage, …)
  - most ontologies VIVO introduce their own “middle concepts” — person, paper, project — these must be “reconciled” with VIVO-ISF. This is fundamental collaborative work across projects to be facilitated by OpenRIF

Hosted by ALCTS, Association for Library Collections and Technical Services.
Karma

A data integration tool

http://usc-is-i2.github.io/karma/

Hosted by ALCTS, Association for Library Collections and Technical Services.
Karma

- Examine your data
- Clean your data
- Create local ontology extensions (optional)
- Model your data
- Load your data

Hosted by ALCTS, Association for Library Collections and Technical Services.
Karma enables integration of:

- CSV/TSV files
- XML
- JSON
- Databases
- KML
- Web APIs

Hosted by ALCTS, Association for Library Collections and Technical Services.
The modeling consists of four steps:

- Assign Semantic Types
- Constructing the Graph
- Refine Source Model
- Generate Formal Specification
Karma import options:

- Database Table
- Using SQL
- From Service
- From File

Hosted by ALCTS, Association for Library Collections and Technical Services.
## Importing CSV/TSV files

<table>
<thead>
<tr>
<th>org_name</th>
<th>org_ID</th>
<th>org_IDurl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounting</td>
<td>8b9df681</td>
<td><a href="http://vivo.northwestern%E2%80%A6">http://vivo.northwestern…</a></td>
</tr>
<tr>
<td>Aerospace Engineering</td>
<td>e5c84f0a</td>
<td><a href="http://vivo.northwestern%E2%80%A6">http://vivo.northwestern…</a></td>
</tr>
<tr>
<td>Agriculture Leadership</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education and Communication</td>
<td>2d953da5</td>
<td><a href="http://vivo.northwestern%E2%80%A6">http://vivo.northwestern…</a></td>
</tr>
</tbody>
</table>

Hosted by ALCTS, Association for Library Collections and Technical Services.
Modeling organizations/units data

Hosted by ALCTS, Association for Library Collections and Technical Services.
R2RML - RDB to RDF Mapping Language

“R2RML, a language for expressing customized mappings from relational databases to RDF datasets. Such mappings provide the ability to view existing relational data in the RDF data model, expressed in a structure and target vocabulary of the mapping author's choice. R2RML mappings are themselves RDF graphs and written down in Turtle syntax. R2RML enables different types of mapping implementations. Processors could, for example, offer a virtual SPARQL endpoint over the mapped relational data, or generate RDF dumps, or offer a Linked Data interface.”

http://www.w3.org/TR/r2rml/
Workbench: organizations/units N-Triples

<table>
<thead>
<tr>
<th>Subject</th>
<th>Predicate</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://vivo.northwestern.edu/n8b9df681">http://vivo.northwestern.edu/n8b9df681</a></td>
<td><a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a></td>
<td><a href="http://vivoweb.org/ontology/core#AcademicDepartment">http://vivoweb.org/ontology/core#AcademicDepartment</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/n8b9df681">http://vivo.northwestern.edu/n8b9df681</a></td>
<td><a href="http://www.w3.org/2000/01/rdf-schema#label">http://www.w3.org/2000/01/rdf-schema#label</a></td>
<td>&quot;Accounting&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/n8b9df681">http://vivo.northwestern.edu/n8b9df681</a></td>
<td><a href="http://vivo.northwestern.edu/ontology/vlocal#orgID">http://vivo.northwestern.edu/ontology/vlocal#orgID</a></td>
<td>&quot;8b9df681&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nt5884f0a">http://vivo.northwestern.edu/nt5884f0a</a></td>
<td><a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a></td>
<td><a href="http://vivoweb.org/ontology/core#AcademicDepartment">http://vivoweb.org/ontology/core#AcademicDepartment</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nt5884f0a">http://vivo.northwestern.edu/nt5884f0a</a></td>
<td><a href="http://www.w3.org/2000/01/rdf-schema#label">http://www.w3.org/2000/01/rdf-schema#label</a></td>
<td>&quot;Aerospace Engineering&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nt5884f0a">http://vivo.northwestern.edu/nt5884f0a</a></td>
<td><a href="http://vivo.northwestern.edu/ontology/vlocal#orgID">http://vivo.northwestern.edu/ontology/vlocal#orgID</a></td>
<td>&quot;nt5884f0a&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nz6q3ydaq">http://vivo.northwestern.edu/nz6q3ydaq</a></td>
<td><a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a></td>
<td><a href="http://vivoweb.org/ontology/core#AcademicDepartment">http://vivoweb.org/ontology/core#AcademicDepartment</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nz6q3ydaq">http://vivo.northwestern.edu/nz6q3ydaq</a></td>
<td><a href="http://www.w3.org/2000/01/rdf-schema#label">http://www.w3.org/2000/01/rdf-schema#label</a></td>
<td>&quot;Agriculture Leadership Education and Communication&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nz6q3ydaq">http://vivo.northwestern.edu/nz6q3ydaq</a></td>
<td><a href="http://vivo.northwestern.edu/ontology/vlocal#orgID">http://vivo.northwestern.edu/ontology/vlocal#orgID</a></td>
<td>&quot;q63y3ydaq&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nz7j3b992">http://vivo.northwestern.edu/nz7j3b992</a></td>
<td><a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a></td>
<td><a href="http://vivoweb.org/ontology/core#AcademicDepartment">http://vivoweb.org/ontology/core#AcademicDepartment</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nz7j3b992">http://vivo.northwestern.edu/nz7j3b992</a></td>
<td><a href="http://www.w3.org/2000/01/rdf-schema#label">http://www.w3.org/2000/01/rdf-schema#label</a></td>
<td>&quot;Animal Science&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nz7j3b992">http://vivo.northwestern.edu/nz7j3b992</a></td>
<td><a href="http://vivo.northwestern.edu/ontology/vlocal#orgID">http://vivo.northwestern.edu/ontology/vlocal#orgID</a></td>
<td>&quot;7j3b992&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nz6dd992">http://vivo.northwestern.edu/nz6dd992</a></td>
<td><a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a></td>
<td><a href="http://vivoweb.org/ontology/core#AcademicDepartment">http://vivoweb.org/ontology/core#AcademicDepartment</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nz6dd992">http://vivo.northwestern.edu/nz6dd992</a></td>
<td><a href="http://www.w3.org/2000/01/rdf-schema#label">http://www.w3.org/2000/01/rdf-schema#label</a></td>
<td>&quot;Anthropology&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nz6dd992">http://vivo.northwestern.edu/nz6dd992</a></td>
<td><a href="http://vivo.northwestern.edu/ontology/vlocal#orgID">http://vivo.northwestern.edu/ontology/vlocal#orgID</a></td>
<td>&quot;6dd992&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nxct22844">http://vivo.northwestern.edu/nxct22844</a></td>
<td><a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a></td>
<td><a href="http://vivoweb.org/ontology/core#AcademicDepartment">http://vivoweb.org/ontology/core#AcademicDepartment</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nxct22844">http://vivo.northwestern.edu/nxct22844</a></td>
<td><a href="http://www.w3.org/2000/01/rdf-schema#label">http://www.w3.org/2000/01/rdf-schema#label</a></td>
<td>&quot;Architecture&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/nxct22844">http://vivo.northwestern.edu/nxct22844</a></td>
<td><a href="http://vivo.northwestern.edu/ontology/vlocal#orgID">http://vivo.northwestern.edu/ontology/vlocal#orgID</a></td>
<td>&quot;ct22844&quot;</td>
</tr>
</tbody>
</table>
Modeling the person file

Hosted by ALCTS, Association for Library Collections and Technical Services.
Workbench: person N-Triples

<table>
<thead>
<tr>
<th>Subject</th>
<th>Predicate</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://vivo.northwestern.edu/individual/2001-01-01T00:00:00">http://vivo.northwestern.edu/individual/2001-01-01T00:00:00</a></td>
<td><a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a></td>
<td><a href="http://vivoweb.org/ontology/core#DateTimeValue">http://vivoweb.org/ontology/core#DateTimeValue</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/individual/2001-01-01T00:00:00">http://vivo.northwestern.edu/individual/2001-01-01T00:00:00</a></td>
<td><a href="http://vivoweb.org/ontology/core#DateTimePrecision">http://vivoweb.org/ontology/core#DateTimePrecision</a></td>
<td><a href="http://vivoweb.org/ontology/core#yearPrecision">http://vivoweb.org/ontology/core#yearPrecision</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/individual/2001-01-01T00:00:00">http://vivo.northwestern.edu/individual/2001-01-01T00:00:00</a></td>
<td><a href="http://vivoweb.org/ontology/core#dateTime">http://vivoweb.org/ontology/core#dateTime</a></td>
<td>&quot;2001-01-01T00:00:00&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/individual/n2de06eq">http://vivo.northwestern.edu/individual/n2de06eq</a></td>
<td><a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a></td>
<td><a href="http://vivoweb.org/ontology/core#FacultyMember">http://vivoweb.org/ontology/core#FacultyMember</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/individual/n2de06eq">http://vivo.northwestern.edu/individual/n2de06eq</a></td>
<td><a href="http://www.w3.org/2000/01/rdf-schema#label">http://www.w3.org/2000/01/rdf-schema#label</a></td>
<td>&quot;Sadav, Manjit S.&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/individual/n2de06eq">http://vivo.northwestern.edu/individual/n2de06eq</a></td>
<td><a href="http://vivoweb.org/ontology/core#hrJobTitle">http://vivoweb.org/ontology/core#hrJobTitle</a></td>
<td>&quot;Professor&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/individual/n2de06eq">http://vivo.northwestern.edu/individual/n2de06eq</a></td>
<td><a href="http://vivoweb.org/ontology/core#end">http://vivoweb.org/ontology/core#end</a></td>
<td><a href="http://vivo.northwestern.edu/individual/2014-01-01T00:00:00">http://vivo.northwestern.edu/individual/2014-01-01T00:00:00</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/individual/n2de06eq">http://vivo.northwestern.edu/individual/n2de06eq</a></td>
<td><a href="http://vivoweb.org/ontology/core#start">http://vivoweb.org/ontology/core#start</a></td>
<td><a href="http://vivo.northwestern.edu/individual/2000-10-01T00:00:00">http://vivo.northwestern.edu/individual/2000-10-01T00:00:00</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/individual/n2de06eq">http://vivo.northwestern.edu/individual/n2de06eq</a></td>
<td><a href="http://vivo.northwestern.edu/ontology/vlocal#netID">http://vivo.northwestern.edu/ontology/vlocal#netID</a></td>
<td>&quot;Sadav&quot;</td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/individual/n2de06eq">http://vivo.northwestern.edu/individual/n2de06eq</a></td>
<td><a href="http://purl.obolibrary.org/obo/ARG_20000028">http://purl.obolibrary.org/obo/ARG_20000028</a></td>
<td><a href="http://vivo.northwestern.edu/individual/nvcard2de06eq">http://vivo.northwestern.edu/individual/nvcard2de06eq</a></td>
</tr>
<tr>
<td><a href="http://vivo.northwestern.edu/individual/nc81ef444">http://vivo.northwestern.edu/individual/nc81ef444</a></td>
<td><a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a></td>
<td><a href="http://vivoweb.org/ontology/core#FacultyMember">http://vivoweb.org/ontology/core#FacultyMember</a></td>
</tr>
</tbody>
</table>

Hosted by ALCTS, Association for Library Collections and Technical Services.
Modeling the position file

Hosted by ALCTS, Association for Library Collections and Technical Services.
Modeling publications data - academic articles
Getting Data from your VIVO: SPARQL
All Things have URI

- All things in VIVO have a Uniform Resource Locator (URI)
- In VIVO, URI usually look like http://yourplace/individual/nxxxxxx
- Violeta Ilik’s URI at OpenVIVO is: http://openvivo.org/a/orcid0000-0003-2588-3084
We want to use VIVO Data

• “Use”
  – Make lists
  – Count things
  – Make reports
  – Get data out of VIVO for use in Excel
    • Or statistics software (R, SAS, SPSS, Stata, …)
    • Or reporting software (Crystal Reports, …)
    • Or visualization software (R, …)

Hosted by ALCTS, Association for Library Collections and Technical Services.
SPARQL queries

- SPARQL is a “query language” for asking for data from a set of triples
- SPARQL is fun and easy. But like any computer language, it is precise, in other words, you will get what you ask for.
Prefixes for the names of ontologies

# List the predicates and objects of all triples for a specified subject

Comments to describe your query

A sample query
Sample SPARQL queries

http://mconlon17.github.io/sparql

SPARQL queries tested with VIVO 1.9 on UF VIVO
Each is marked with SPARQL level of difficulty
Reports, Data Management, People, Papers, Grants, Organizations
Real-world examples. Most are the result of questions asked by stakeholders

Hosted by ALCTS, Association for Library Collections and Technical Services.
Acknowledgments:

Paul Albert - Weill Cornell Medicine
Jim Blake - Cornell University
Michael Conlon - VIVO Project Director
Jon Corson-Rikert
John Fereira - Cornell University
Huda Kahn - Cornell University
Muhammad Javed - Cornell University
Julia Trimmer - Duke University

Hosted by ALCTS, Association for Library Collections and Technical Services.
Thank you

Violeta Ilik

https://vioil.github.io/
http://openvivo.org/a/orcid0000-0003-2588-3084

Hosted by ALCTS, Association for Library Collections and Technical Services.