# VIVO and the role of librarians part 1

Violeta Ilik

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



### 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

# Goals for today - 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.

### Goals for 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

# A brief history

2003-2005 Cornell created VIVO as a relational database for the life sciences

2006-2008 Cornell expanded to all disciplines and converted VIVO to Semantic Web application
2009-2012 VIVO was awarded a NIH grant
2013-2015 VIVO joins DuraSpace open-source portfolio preserving digital scholarship

# VIVO

An *open-source semantic web application* that enables the discovery of research and scholarship across disciplines in an institution.

VIVO harvests data from *verified* sources and offers *detailed profiles* of faculty and researchers.

*Public, structured linked data* about investigators interests, activities and accomplishments, and *tools to use that data to advance science*.

VIVO enjoys a *robust open community space* to support implementation, adoption, & development efforts around the world. See *http://wiki.duraspace.org/display/VIVO* 

### 

Iome People Organizations Research Events



# VIVO is open

- Open source, open community, open data
- Flexible and customizable
- Enterprise application
- Authoritative, verified data



# VIVO is about researchers

### • Expert locator

### • Public-facing web profiles





# Why do we want VIVO?



# Integrating institutional data



VIVO

# VIVO data is visualized

### Browse global collaborations

(2012 - 2017)

Filter by: Academic Unit

Filter by: Subject Area

Filter by: Publication Year

#### Clear Filter



### Global Collaborations

#### .



6

Export as JSON

Show table format

Export as SVG

# VIVO data is visualized



VIVO

# VIVO data is visualized



VIVO

# VIVO data is visualized

#### Browse publication venues

Select: Journal or Proceedings

	Search
	Proceedings of the National Academy of Sciences of the United States of America (508)
	Applied Physics Letters (408)
	Journal of Dairy Science (353)
	Journal of Geophysical Research (293)
	Science (New York, N.Y.) (292)
	Physical Review Letters (274)
	PLoS One 6 (265)
	Nature (242)
	Plant Physiology (239)

 Some publications are co-authored by scholars who are in multiple academic units.

#### Journal/Proceedings Insights a Show table format Now showing: Nature Export as JSON Export as SVG Click on a bar to view the list of publications for an academic unit. Click on an academic unit name to filter the data in the view ount 20 ation College of Agriculture and Life Sciences 18 16 -1. Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium oxysporum 14-2. Differential innate immune signalling via Ca<sup>2+</sup> sensor protein kinases 12 -3. Evolutionary biology: Expanding islands of speciation (News and Views article) 10 -4. Genome sequencing and analysis of the model grass Brachypodium distachyon 8 -5. Integrating common and rare genetic variation in diverse human populations 6 -6. Monitoring the world's agriculture 4 . Publication Year College of Agriculture and Life Sciences

College of Human Ecology College of Veterinary Medicine Computing and Information Science

# VIVO data is linked

Sinnott-Armstrong, W. "Explanation and justification in moral epistemology." 1999. Link to Item

#### **Digital Publications**

Sinnott-Armstrong, W. <u>"My Brain Made Me Do It — So What?."</u> (March 2015). Link to Nam

Sinnott-Armstrong, W. <u>"Does Philosophy Matter?."</u> (March 2015 Link to Itam

Sinnott-Armstrong, W. "My Brain Made Me Do It, but Does that Matter?." (December 2014). Link to Item

Sinnott-Armstrong, W. "Moral Skepticism." (2004). Link to Item

Sinnott-Armstrong, W. "Consequentialism." (2003). Link to Item

Show More Digital Publications.

#### Journal Articles

Clifford, S, Iyengar, V, Cabeza, R, and Sinnott-Armstrong, W. "Moral foundations vignettes: a standardized stimulus database of scenarios based on moral foundations theory." Behavior research methods 47, no. 4 (December 2015): 1178-1198. FullTast

Summers, JS, and Sinnott-Armstrong, W. <u>"Scrupulous agents."</u> *Philosophical Psychology* 28, no. 7 (October 3, 2015): 947–966.

Schlegel, A, Alexander, P, Sinnott-Armstrong, W, Roskies, A, Tse, PU, and Wheatley, T. <u>"Hypnotizing Libet: Readiness potentials</u> <u>with non-conscious volition."</u> *Consciousness and cognition* 33 (May 2015): 196-203. putter.

# Automatically links people to pubs, grants, collaborators--any data element

#### Home People Schools / Institutes > Research About

### My Brain Made Me Do It — So What?

Digital Publication

Full Text

Link to Item

**Duke Authors** 

Sinnott-Armstrong, Walter

**Cited Authors** 

Sinnott-Armstrong, W

**Published Date** 

March 2015

Published By

Oxford Uehiro Center for Practical Ethics

# VIVO can:

- Include researchers in all disciplines
- Be internationalized in many languages
- Consume data from ORCiD, Figshare, ...
- Be used on mobile devices



# Extended search



# Site tailored to finding experts

### https://scholars.uow.edu.au/

#### **Capability Map** Build a 'first pass' capability map by typing in a search term that could be said to represent a broad research capability. Linked data Cutoff: 10 Search and Expand Reset Search Search terms Info pause hide group labels Open source software Current search terms Semantic Integration (Computer systems) Delete selected PL/SQL (Computer program language) Da'sQL\*PLUS'(Computer pr Coffee growers; social condition Friedman Artificial Intelligence Social media puter program language)--Computer programs larve SQL (Computer program language)--Standards Java Serint (Com im language SQL/ORACLE (Computer program language Computational linguistics Database design Data dictionarias Enterprise Resource Planning gretry-Environmental aspects logy-Research SPARQL (Computer program language) Institutional repositories Comontic We RDF (Document markup language Database security database systems Open source software Chemistr Libraries Database selection 2 Sustainability information science Big data Metadata Data mining tion storage Information science-Informechamistry forganic -Diplomata--History Catholics, English--History Data curation Bigital humanities centers les-Digital libraries History--Sources Chase Information storage and retrieval systems ----Database design World Wide Web Digital libraries Linked data Humanities--Digital libraries man l'Prosopographymodern, 1500-1700-- Societies, etc. linked data Big data Information technology--Research Image processing Science publishing Bioinformatica Riginformatics Data curation information networks nal lingulatics Federer Alage--Ecology Architecture--Data proceesing Algae-Ecophysiology Capability size indicates resarchers in field rmation technology--Technological Innovations Architectural practice--Data processing Data mining Biological Oceanography Edge color indicates people connected Marine since Obveloing Geoinformatice Data transmission systems--Reliability Information technology; Evaluation oceasing Geospatial data Toxic marine algae Group color indicates degree -Ecology Ecology Databases Cataloging-Data processing + Standardshnology-Data processing achine learni Algae--Growth--Measurement label shows people in group Distoma-Ecoinformatics Information technology--Costs Information storage a Cataloging-Data processing logy>comp -assisted instruction Marine plankton--Ecology 3 links ine ecology v-Research ormation technology--Planning Marcie algae--Ecology ≥4 links Marine Coastal ecology Cillata; Ecology and culture media

Extended capability map in OpenVIVO

### http://openvivo.org/



# "Is VIVO for you?"





# Planning a VIVO implementation

https://wiki.duraspace.org/display/VIVO/Planning+a+VIVO+Implementation

# VIVO: The Big Picture



# VIVO Implementation includes roles:

### Project Management

Managing a VIVO project can require everything from finding resources to facilitating the structure, mission and strategies.

- Outreach and Community Engagement Your VIVO community includes many stakeholders and sponsors from across the institution -- communications are critical.
- Data Management

VIVO is all about the data: where to find it, how to load it, and how to keep it as updated as possible.

### Technical Development

Developers, programmers, system admins, and other technical folks - System Architecture, Identify Customizations, Establish Data Feeds, Develop Prototypes, Build Customized System, Test Performance, Provide System Support, Implement System Upgrades



# VIVO Implementation includes phases:

- Analysis
- Design
- Implementation
- Launch
- Maintenance

### • Project Management

Analysis: Establish Governance; Resource Identification Design: Branding, Further Define Scope, Request data Feeds Implementation: Create Launch Strategy Launch: Oversee Publicity Campaign, Implement Assessment Plan Maintenance: Contribute to VIVO community

Outreach and Community Engagement

Analysis: Identify Stakeholders, Gather Use Cases Design: Share Prototypes and/or existing VIVOs: <u>http://duraspace.org/registry/vivo</u> Implementation: Identify Power Users, Develop Training Materials Launch: Publicize VIVO, Hold Training Sessions Maintenance: Find New Collaborators, Hold User Meetings

### Data Management

Analysis: Identify Potential Data Sources Design: Map Data to Ontologies, Document Data Cleanup Strategy Implementation: Prepare Data Loads, Document Data Provenance Launch: Route Data Cleanup Requests, Support Data Provisioning Maintenance: Manage Ontology Updates, Add New Data & Sources

### Technical Development

Analysis: Learn System Architecture, Identify Customizations Design: Establish Data Feeds, Develop Prototypes Implementation: Build Customized System, Test Performance Launch: Provide System Support, Implement System Upgrades Maintenance: Develop New Features

## **VIVO success indicators**

- High-level sponsorship
- Commitment to ongoing support
- Alignment with institutional plan

# It's all about goals

What's most important to your institution:

- Public-facing web profiles?
- Research discovery?
- Platform for open science?
- Others?



Your goals must inform the strategy

# **Rollout** strategies

- Broad and shallow?
  - Lots of people, less data
  - Add data over time
- Narrow and deep?
  - Fewer people, more data
  - Add people over time





# Example 1

Goal: Public profiles for 3,000 faculty members Strategy: Implement "extended directory" first and add data in phases

# Example 2

Goal: Show collaborations within a research group Strategy: Create profiles for group members with publications, grants, research interests, keywords

# Example 3

### Goal: Encourage users to improve data

Strategy: Introduce editing process in private profiles; communicate plans for going live





# Assembling a project team

### Technical:

- Programming lead (0.5 1 FTE)
- Programmers/developers (1 3 FTEs/2 bodies)
- Database administrators
- Systems administrators (networks, servers)

# Assembling a project team

Project/business:

- Project or product manager (0.5 1 FTE)
- Data analyst or curator, plus ontologist (0.5 1 FTE)
- Communicator or trainer or outreach person (0.25 – 1 FTE)

# **Other allies**

- Librarians, scholarly communications officers
- Chief academic officer (faculty titles, appointments)
- Grants offices, other data stewards
- News office, communicators
- Web designers, corporate identity
- Anyone who cares

# **Budgeting & funding**

- Costs involved IT infrastructure and human resources
- Always estimate pessimistically
- Funding model can evolve (starter funding now, ongoing funding later)
- Provost, library split
- Or contributions by % of faculty organization
- Budget for DuraSpace membership

# **Communication strategy**

- Big implementation?
  - Use institutional resources
  - Keep message simple
- Smaller group?
  - Use local resources
  - More events
  - More complex information



# You need a plan

- Strategy, timeline, goals, tactics, resources
- Tactics: home page, email lists, Twitter, news stories, promotional videos
- Events are important:
  - Demos
  - Development events
  - Attend faculty meetings

# **Best practices for marketing**

- Focus on specific benefits to each stakeholder
- Repetition is important
- Keep emails short with links to more info
- Be positive and patient
- Facilitate community engagement, conversations

# Things you may hear

- Faculty don't like to be surprised.
- Don't notify faculty until it is ready.
- Faculty won't come to training or demos.
- "How was I supposed to know about this?"
- "I don't read email."
- "I don't read instructions."

### How to Pitch your VIVO



# **Tailoring your VIVO Pitch**

- > How do you explain VIVO?
- Complex, multi-layered system
- Different people care about different aspects
- > Don't bury people with technical details

# **Summary**

- > Identify organizational goals
- Choose a stakeholder group
- > List unique, strategic benefits
- > Create pitch:
  - short
  - sweet



# **Best practices for support:**

- ★ Create a local network
- ★ Provide documentation, videos
- ★ Enable help ticket submission
- ★ Give ongoing "info sessions"
- $\star$  Hold events and user meetings



# Duke's Support Model



# 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





### Thank you See you on March 14th VIVO and the role of librarians part 2

Violeta Ilik

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

