Introduction to VIVO VIVO 2014 Conference Workshop

Jon Corson-Rikert | Lamont Cannon | Kristi Holmes | Alex Viggio



Overview

- Introductions all around
- Why VIVO?
- VIVO as a semantic web application
- Vitro as a platform
- The VIVO community
- Case study vignettes
- Planning and implementation
- Joining an open source community project



Instructors

- Jon Corson-Rikert
 - VIVO Core Development WG lead
 - Head of Mann Library Information Technology Services, Cornell University Library
- Lamont Cannon
 - Outreach Coordinator for Faculty Data, Office of the Provost, Duke University
- Kristi Holmes
 - VIVO Engagement WG lead
 - Director, Galter Health Sciences Library, Northwestern University
- Alex Viggio
 - VIVO Implementation WG lead
 - Senior Technical Consultant, Symplectic, Ltd.



Why VIVO?



The value of scientific knowledge dispersed across the world can increasingly be captured by those who build networks to take the local to global scale and bring the global back for local impact.

> From "The United States Looks to the Global Science, Technology, and Innovation Horizon," E. William Colglazier and Elizabeth E. Lyons, citing The New Invisible College, Caroline S. Wagner, The Brookings Institution Press, 2008

The research landscape

- Cross–disciplinary
- Multi-institutional
- Organizationally virtual
- Collaborative
- Competitive

Data, information, ideas, and technologies do not propagate within social networks simply on their own merits.

> From "Enabling Scientific Collaboration and Discovery through Semantic Connection," research proposal to NSF. Matthew Mayernik et. al., University Corporation for Atmospheric Research

Latent connectivity

- Social connections fuel collaboration
- Many are not visible
- Finding additional potential connections is beneficial to research
- Transparency helps the public and policymakers as well



Additional incentives

- OSTP mandate

 Goes beyond publications
 Emphasis on data
 Obligation to maintain public access

 Trends in funding opportunities
 ORCID
- SciENCV

Technology landscape

- Mobile platforms
- Lightweight interactive apps
- Visualization libraries
- Search technologies
- Multiplicity of collaboration platforms
- Hunger for alternatives to email, wikis, and websites



Strategic opportunities

- Current Research Information Systems (CRIS) and research networking emerging from the early adopter phase
- Sense of urgency among funders
- Broader community of linked data – E.g., traction in libraries
- Global interest in entity resolution

What is VIVO?



VIVO

A semantic-web-based research and researcher discovery tool

- People plus information on the research they do

- Publicly-visible information, across disciplines
 For external as well as internal audiences
- An open, shared platform for connecting scholars, research communities, campuses, and the world using Linked Open Data (LOD)

A brief VIVO history to the present

2003-2005 First realization for the life sciences at Cornell, as a relational database Expansion to all disciplines at Cornell, 2006-2008 and conversion to Semantic Web 2009-2012 National Institutes of Health-sponsored VIVO: Enabling the National Networking of Scientists project transforms VIVO to a multi-institutional open source platform 2013-2014 VIVO incubator project with DuraSpace for open community development

Key VIVO principles

- Open software
- Open data
- Open ontology
- Open community
- Local control



What VIVO does

- Integrates multiple sources of data
- Provides a content review and editing interface
- Provides a uniform, research-focused institutional web presence for search and browse
- Provides filterable feeds to other websites
- Serves local data to the global linked data cloud

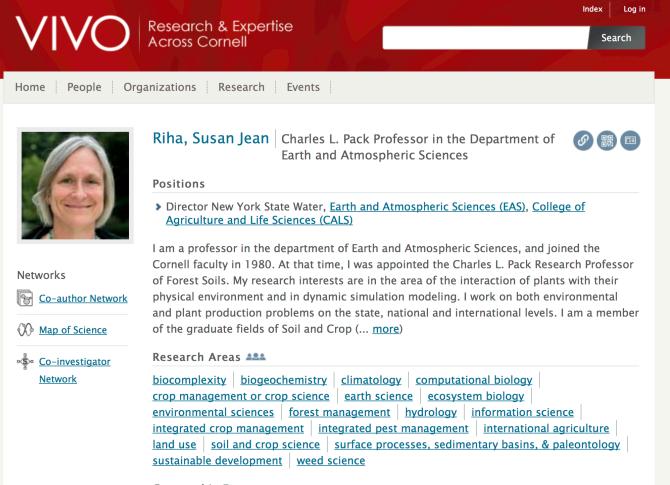
What VIVO models

People and more

- Organizations, grants, programs, projects, publications, datasets, events, facilities, and research resources
- Relationships among the above
 - Meaningfully typed
 - Bi-directionally navigable
- Links to global URIs
 - Vocabularies, ORCID iDs, DOIs, other identifiers for people, places, organizations, events



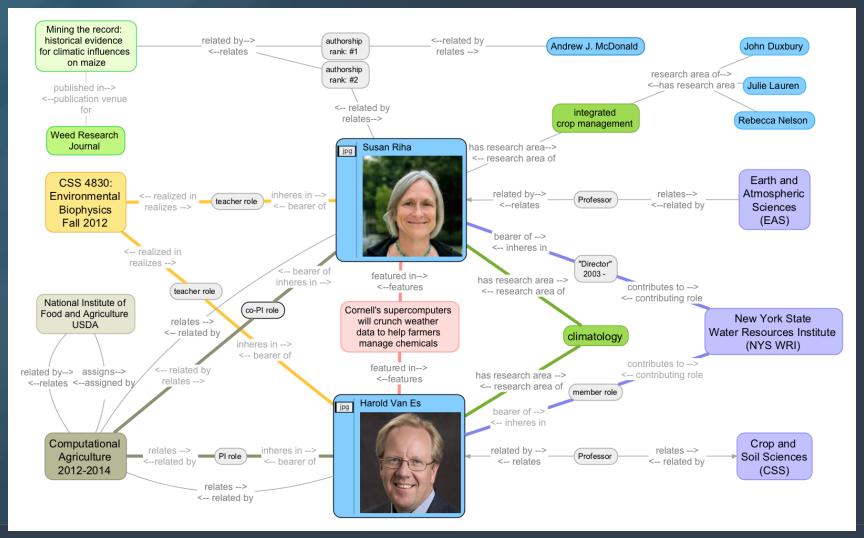
People



Geographic Focus



And how they connect



VIVO

Co-Author Network (GraphML File)

Profile

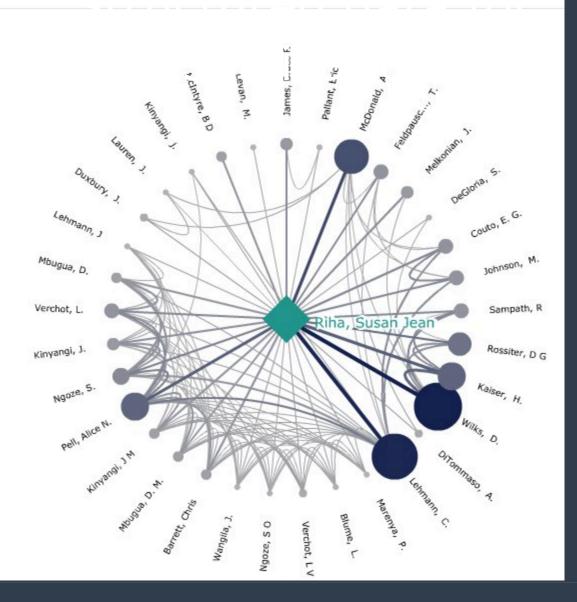


Riha, Susan Jean Charles L. Pack Professor in t...

VIVO profile | Co-author network

- 132 Publication(s)
- 33 Co-author(s)
- 1980 First Publication
- 2010 Last Publication

Note: This information is based solely on publications which have been loaded into the VIVO system. This may only be a small sample of the person's total work.



Typical data sources

- HR people and their positions
- Research administration grants & contracts
- Registrar courses
- Faculty reporting awards, professional service, education, research areas, geographic focus
- External repositories Pubmed, Scopus, Web of Science, CrossRef
- Events calendar
- News -- internal and external
- Extension outreach, technology transfer



Value for institutions

- Common data substrate serving many functions
 - Cross-silo
 - Granular
 - Public-facing
- Distributed data entry and curation
 - Editing by proxy or through data relationships
 - Filling gaps between systems of record
 - Directing changes back to systems of record
- Direct linking across campuses
- Data sharable in a standard format

Data ecosystem stewardship

- Manage data at its appropriate source with appropriate privacy
 - HR, grants management, registrar, graduate school, colleges and schools, research centers, extension
 - Department/agency/division/geographic location/ research unit
- Consciously derive public data for exchange
- Engage stakeholders and build relationships
- Recruit power users for training and local knowledge
- Data that are visible get corrected!

Enter data once, use it many times



Home People

Contact information

Positions O

Web Pages 🕒

- Abruña Group Chemistry and Chemical Biology profile | D
- Preferred Title Emile M. Chamot Professor Chemistry and Chemical Bic 2 🏛 Overview

Research & Expertise Across Cornell

Resource URI: http://vivo.cornell.ed Abruña. Héctor D

Edit this individual

Organizations Research

Admin Panel

The Abruña Group focuses on a wide variety of techniques for molecular electronics.

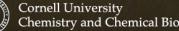
Research Areas O

| Affiliation | Research | Publications | Teaching |
|-------------|----------|---------------------|-----------------|
| Affiliat | ion | | |

- head of G
- Cornell Fuel Cell Institute (CFCI)



Index Site Admin Jon



Chemistry and Chemical Biology

Undergraduate Research Faculty

Abruña, Héctor D

E. M. Chamot Professor



email: hda1@cornell.edu phone: 607-255-4720 room: Olin Chemistry Research Wing

Websites

Abruña Group

Department Appointments

Chemistry and Chemical Biology (CHEM)

Graduate Fields

- Chemistry and Chemical Biology
- Other Affiliations

Overview

Courses

You are here: Chemistry and Chemical Biology > Faculty > Faculty Detail

Graduate

The Abruña Group focuses on the development and characterization of new materials using a wide variety of techniques for fuel cells, batteries, and molecular assemblies for molecular electronics.

Search

Chemistry and Chemical Biology
Cornell

Events and News Directories

Research

Our research effort takes an interdisciplinary approach to the study of electrochemical phenomena. We employ electrochemical techniques as probes of a variety of chemical systems, and we use other techniques such as x-ray based methods, differential electrochemical mass spectrometry, in-situ FT-IR, scanned probe microscopies, scanning electrochemical microscopy, low temperature conductance and spectroscopic techniques to address problems of electrochemical interest. Current areas of research include:

1. Fuel cells:

. The use of ordered intermetallics, such as BiPt for the electrocatalytic oxidation of formic acid, methanol, ethanol and other small organic molecules of potential utility as fuels in fuel cells.

- Use of Differential Electrochemical Mass Spectrometry (DEMS), in-situ FT-IR in for mechanistic studies related to fuel cells.
- Development of in-situ TEM techniques for the study of fuel cell and battery materials
- 2. Electrical Energy Storage (EES): Batteries and Supercapacitors
- Computational screening synthesis and characterization of organic molecules for EES
- In-situ testing of battery systems using in-situ x-ray based technique (XRD, EXAFS, XANES)
- Lithium/sulfur batteries



| Home | People | Organizations | Research | Events | |
|------|--------|---------------|----------|--------|--|
|------|--------|---------------|----------|--------|--|



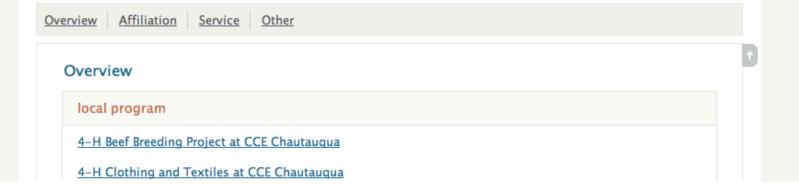
Cornell Cooperative Extension Association of Chautauqua County Cornell Cooperative Extension Association &

Index

Log in

Search





CALS Research and Impact

Information about CALS research projects and their impact throughout the world

HOME

PROJECTS PROJECT LOCATIONS

Search projects...

Q

154 projects

Increasing Resilience to Climate Change in Agriculture in the Middle East

2011 to 2013

Recent Assessment Reports of the Intergovernmental Panel on Climate Change (IPCC) have stated that "warming of the climate system is unequivocal" and have documented widespread evidence of global warming and other major climatic changes, as well as their impacts. These impacts in rural areas – on agriculture, farmers, rural households and rural communities – can be expected to be increasingly substantial. This project reports the results of research to identify and prioritize stakeholder-driven, locally relevant response options to climate change affecting agriculture in the Middle East.

Climate change impacts on Northeast agriculture

2007

Prior climate change studies have focused on major world food crops such as wheat and maize. But this is not relevant to the Northeastern U.S. (NE) agriculture economy, which is dominated by dairy and high-value horticultural crops.

Climate change information for assessing impacts and adaptations

2007 to 2012

Temperature, precipitation, and other climate patterns have changed across the Northeast over the last 30 years, and these trends are expected to continue over the next several decades. Information on the magnitude, direction, and consequences of such trends is essential for farmers, water resource managers, environmental interests, and policy makers in the region. Through this project, we have been able to apply both historical and projected future climate conditions to climate-related decisions.

Climate change and its impact on the distribution of invasive weeds

| Search | |
|-------------------|--|
| climate change | |
| Projects per page | |
| Apply Reset | |

Printer-friendly version PDF version

Filter by Year

| 2020 (4) | |
|----------|--|
| 2019 (4) | |
| 2018 (4) | |
| 2017 (6) | |

2016 (9)

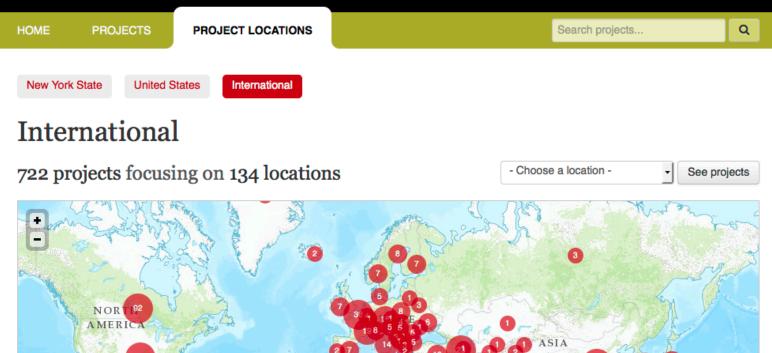
Show more

Filter by New York State County

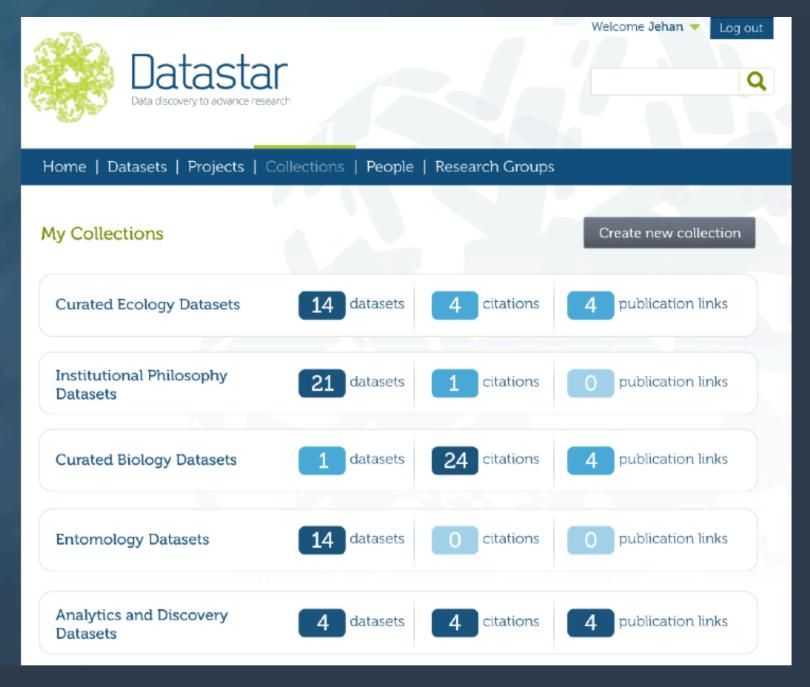
- Tompkins (27)
- Cayuga (24)
- Ontario (16)
- Niagara (15)
- Schuyler (15)
- Show more

CALS Research and Impact

Information about CALS research projects and their impact throughout the world







, VIVC



Welcome Jehan 👻 🛛 Log ou

О

Home | Datasets | Projects | Collections | People | Research Groups

Archaeological Findings | All datasets

Cited Linked to 23 Part of 2 projects 9 collections All links All related datasets

 Identifier
 DOI: 123234

 Authors
 Hardy, Thomas Jehan Sorour

 Contributors
 Dickens, Charles

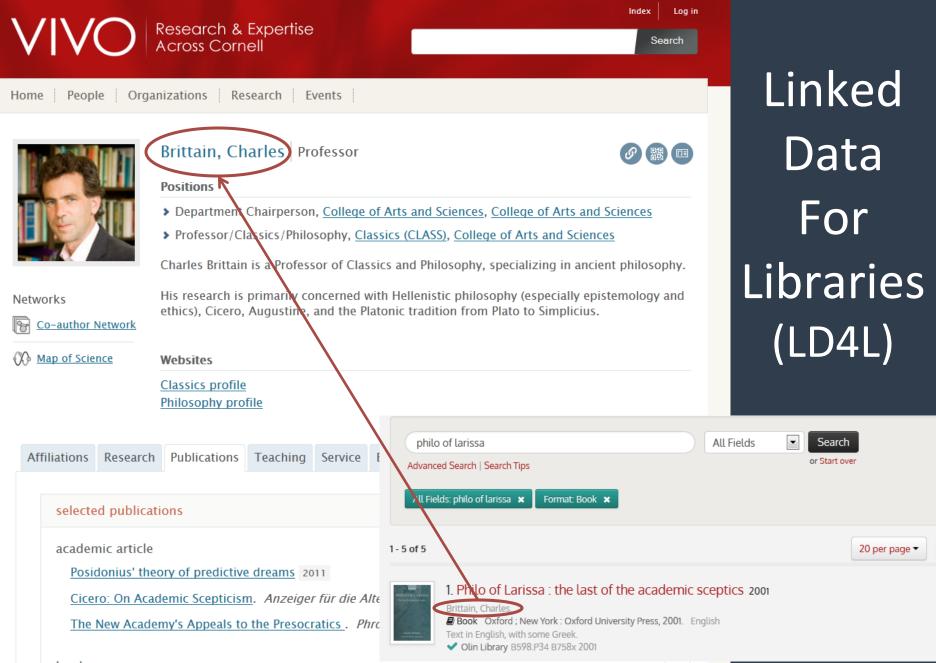
 Description
 Description lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivarnus viverra commodo purus sed euismod. Vestibulum volutpat pellentesque mauris, quis fringilla augue feugiat eget. Nulla quis nibh ac ligula condimentum mollis. Ut sit amet arcu diam, vel omare dui. Aliquam vestibulum sodales mi non ultrices. Morbi tristique laoreet imperdiet. Proin

Cite this dataset: Why Cite?

Thomas, Hardy, URL: www.urlsample.com

Rights and Restrictions: Rights and restrictions content.

11 VIVO



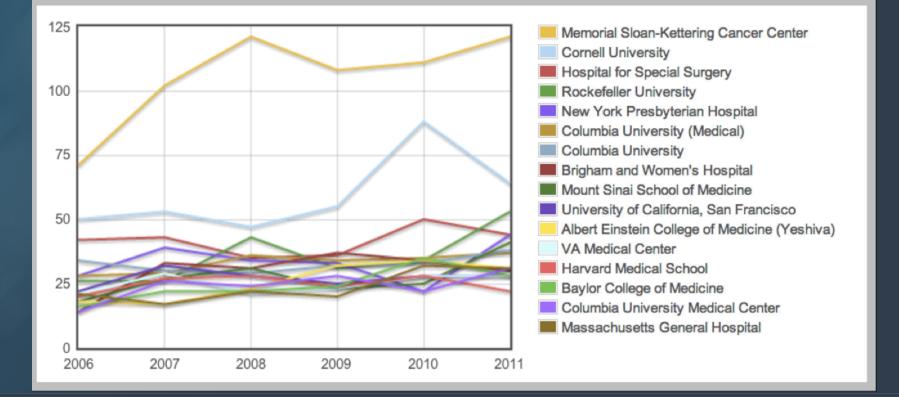
book

Philo of Larissa: The Last of the Academic Sceptics. Oxford ; New York: Oxford University Press.

2001

Weill Cornell data dashboard

• Publication co-authorships by institution





✓ Otolaryngology

Driving questions - Weill Cornell

- Publications with a Weill researcher as first or last author
- PIs having the most collaborations based on grant support
- Publications co-authored with other institutions
- Papers having the most incoming citations
- Researchers publishing the most research articles within a given set of journals in the past 5 years
- Papers receiving federal funding but not deposited in PubMed Central
- Evidence for the institution's open access key opinion leaders



Policy issues

- Dirty data
- Lack even of common definitions of organizational structure or who's faculty
- Data ownership
- Opt-in vs. opt-out and the many dimensions of privacy
- Short-term "go it alone" vs. common good

VIVO and Linked Open Data



The Semantic Web

- Turn data into a web of simple links
- Use ontology to explain how things are linked
- Use reasoning to categorize and add new statements automatically
- Stay flexible and extensible

What is Linked Open Data?

Data

- Structured information, not just documents and text
 A common, simple format
- Open
 - Available, visible, mine-able
 - Anyone can post, consume, and reuse

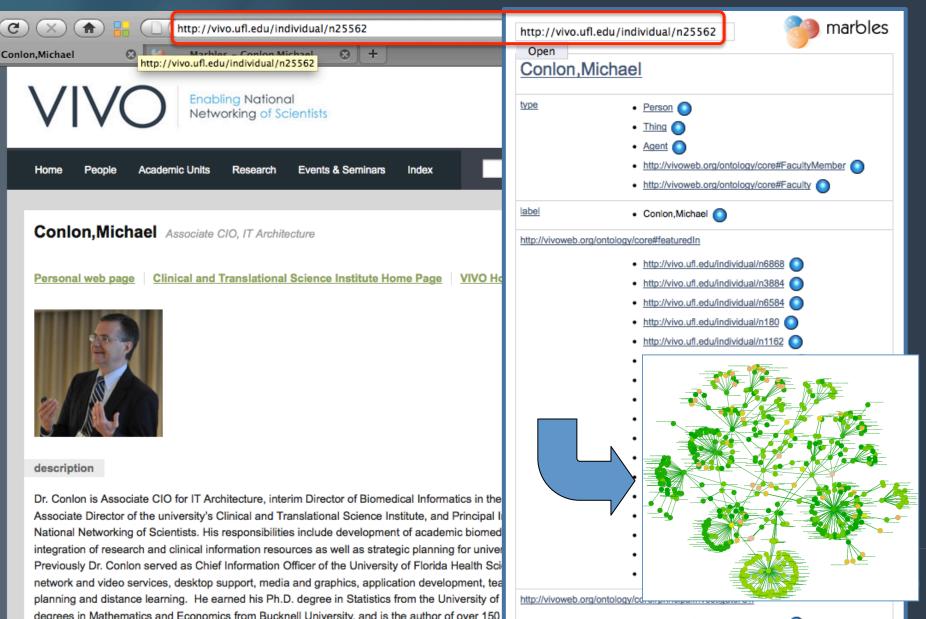
Linked

- Directly by reference

- Indirectly via common references and inference



Linked Open Data



Linked data in OpenAGRIS

v Source:

Centralna Biblioteka Rolnicza/Central Agricultural Library

CBR is a scientific library subordinated to the Ministry of Agriculture and Rural Development. It has branch in Pulawy. CBR collections thematically restricted to agriculture, food processing industry and rel [...]

HOMEPAGE: http://www.cbr.edu.pl/eng/index.php

COVERAGE: Europe

Data from World Bank (double-click an area to zoom)



Thunnus obesus distribution map. Data from Global Biodiversity Information Facility (GBIF)



Thunnus obesus - Global Capture Production (FAO Fishery Statistic)

500

Length-frequency compositions and weight-length relations for bigeye tuna, yellowfin tuna, and albacore (Perciformes: Scombrinae) in the Atlantic, Indian, and eastern Pacific oceans [2008]

RDF dlod live

```
Zhou, Y.
```

```
Zhu, G.
```

Dai, X., Tuna Fishery Technical Working Group of China, Shanghai, China Xu, L., Shanghai Ocean University, Shanghai (China). College of Marine Sciences

Abstract:

Bigeye tuna, Thunnus obesus (Lowe, 1839), yellowfin tuna, Thunnus albacares (Bonnaterre, 1788), and albacore, Thunnus alalunga (Bonnaterre, 1788), are very important species for world fisheries. The weightlength relations (WLRs) of the three species were studied using commonly accepted methodology. Significant differences can be found from the fork length distributions and the WLRs of the above 3 tuna species and the relations of gilled-gutted and whole weight of bigeye and yellowfin tunas collected from the Atlantic, Indian, and Eastern Pacific Oceans. Significant differences of fork length distributions can be found for bigeye tuna, yellowfin tuna, and albacore from the three areas. The date collected will be useful for the fisheries management of the three species studied

Read the article: http://www.aiep.pl/

र्**र +1** { 0

| Agrovoc Keywords: | ▼ Acta Ichthyologica et Piscatoria (Journal) |
|--|---|
| Thunnus obesus body weight Atlantic Ocean fishery data statistical data | FREQUENCY: Semiannual (2 numbers a year) START DATE: 1972 |
| Animal growth forms | Agris articles from the same journal: |
| Tuna Indian Ocean Indian Ocean Thunnus albacares Thunnus albalunga Animal physiology Pacific Ocean Thunnus Fishery production body measurements Fishery management Animal developmental stages | On the occurence of Salomon (Salmo salar L) in the Szczecin Firth and the Lower Odra in 1977 [Poland]. Wplyw detergentu DBS na aktywnosc niektorych enzymow mozgu, skrzeli i surowicy narybku karpia (Cyprinus carpio L). Attainment of sexual maturity by hybrids of rudd, Scardinius erythrophthalmus (L) and carp bream, Abramis brama (L) under experimental conditions No relationship between fecundity and annual |
| | reproductive rate in bony fish |

Data from www.nature.com

- Climatology: Extremes in the Indian Ocean
- Marine biogeochemistry: The ups and downs of ocean oxygen
- Earth science: Subtle minds and mid-ocean ridges
- Ocean-atmosphere coupling: Mesoscale eddy effects

Data from DBPedia

| • | Body weight |
|---|----------------|
| • | Atlantic ocean |
| • | Tuna |
| · | Indian ocean |
| | Thunnus |
| • | Pacific ocean |

http://agris.fao.org/openagris/search.do?recordID=PL2009000495

OpenAgris search augmentation

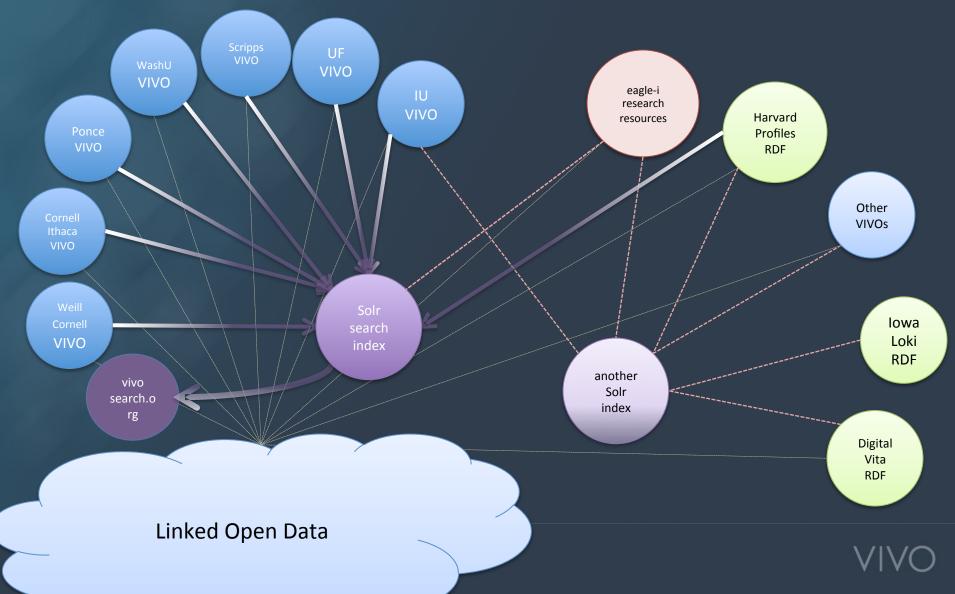
| prove more specific properties of the properise of the properties of the properties of the properti | Find resources | AGRIS About Feedback | |
|--|---|---|--|
| Main based orcopping systems for usualizable agriculture in semi-arid areas of Ethiopia. Passed use based reactions are stream of the system and base of thoses. To alreade this problem the possibility of intercopping, reise y coopping or aller as well of the system and as and of thoses. To alreade this problem the possibility of intercopping, reise y coopping or aller as well of the system and as and of thoses. To alreade this problem the possibility of intercopping, reise y coopping or aller as well of the system and as and of thoses. To alreade this problem the possibility of intercopping, reise y coopping or aller as well of thoses. To alreade the system of th | earch Results ary: maize builts 1 - 10 of 68,194 | | Sort by:) Relevance) Submission Date Order: |
| In character subscience, agroups used and proportidate copping updates and and proportidate copping of makes with efficient legumes, and development of appropriate cope relation systems for hepproved makes production in the d In cARIBS collection since: 1997 Papeopara Intercopping In makes based copping y statems Papeopara Intercopping In makes based copping y statems In cARIBS collection since: 1997 CharaBS collection since: 1994 CharaBS collection since: 2017 CharaBS collection since: 2017 CharaBS collection since: 2017 CharaBS collection since: 2017 CharaBS collection since: 2018 CharBS collection since: 2018< | Maize based cropping systems for sustainable agricult | ure in semi-arid areas of Ethiopia | Refine Search |
| Karono, S.; Flydd, C. (Bala Panellian Tanaman Pangan Malang (Indonesia)) In Pooligoga area the predominant make based "taga" (by lund) ecoping system is a miry season naise monocorp followed by a misseliab-bib bean intercop. The make in the materiab base intercop is nareased abordy affet the od of the millible mainsesses. The babe babe panel preduction at more based abordy affet the od of the millible mainsesses. The babe babe panel preduction at more based abord affet the od of the millible mainsesses. The babe babe panel preduction at more based abord affet the od of the millible mainsesses. The babe babe panel preduction at more based abord affet the od of the millible and preduction at more based abord affet the od of the millible and preduction at more based abord affet the od of the millible and preduction at more based abord affet the od of the millible and more babe babe panel preduction. The millible and material base calculates abord at the odd of the millible and preduction. The millible and material base calculates abord at the odd of the millible and material base calculates abord at the odd of the millible and material base calculates abord at the odd of the millible and material for the odd of the millible and the base and material for the odd of the millible and the millible and material for the odd of the millible and for the millible and material for the odd of the millible and the material for the odd of the millible and the abord and for the material for the odd of the millible and the abord and material for the odd of the millible and the odd of the millible and the abord and the millible and the abord and | In peasant subsistence agriculture use of appropriate cropping syste production was identified as one of the major production constraints alley cropping of maize with efficient legumes, and development of a | ms enable the farmers to use natural resources efficiently. Lack of appropriate cropping systems to suit maize in the semi-arid areas of Ethiopia. To alleviate this problem the possibility of intercropping, relay cropping or | zea mays (24442) maire (11422) crop yield (4069) varieties (2445) hybrids (2373) yields (2280) growth (1710) |
| Breeding of speciality maize for industrial purposes Language: sr Paice, A, Maize Research Institute Zamun Polip, Belgrade - Zamun (Serbia); Paiovic, M., Maize Research Institute Zamun Polip, Belgrade - Zamun (Serbia); Paiovic, M., Maize Research Institute Zamun Polip, Belgrade - Zamun (Serbia); Paivox, M., Maize Research Institute Zamun Polip, Belgrade - Zamun (Serbia); Paivox, M., Maize Research Institute Zamun Polip, Belgrade - Zamun (Serbia); Paivox, M., Maize Research Institute Zamun Polip, Belgrade - Zamun (Serbia); Paivox, M., Maize Research Institute Zamun Polip, Belgrade - Zamun (Serbia); Side, J., Maize Image: Strength | Karsono, S.; Flyod, C. (Balai Penelitian Tanaman Pangan Malang (II In Probolinggo area the predominant maize based "tegal" (dry land) the maize/lab-lab intercop is harvested shortly after the end of the r the variable late rains. Pigeonpea could be used in a similar way to l | ndonesia)) cropping system is a rainy season maize monocrop followed by a maize/lab-lab bean intercrop. The maize in elable rainy season. The lab-lab bean grows through into the dry season, utilising residual soil moisture and | Genotype and environmental interaction effect heterosis expression in maize |
| Determination of plant population and planting time in maize (Zea mays L) and climbing bean (Phaseolus vulgaris L) intercropping system Breeding of maize types with specific traits Negash Geleta (Bako Agricultural Research Center, Bako (Ethiopia)); Chemeda Daba (Bako Agricultural Research Center, Bako (Ethiopia)); Setegn Gebeyehu (Bako Agricultural Breeding of maize types with specific traits Negash Geleta (Bako Agricultural Research Center, Bako (Ethiopia)); Chemeda Daba (Bako Agricultural Research Center, Bako (Ethiopia)); Setegn Gebeyehu (Bako Agricultural Breeding of maize types with specific traits Names Serveriment was conducted at Bako research center in 1998 and 1999 cropping seasons. The objectives of the study were to determine agronomically optimum plant Maize rough dwarf - Maize rough dwarf fijk Biotechnology in maize bard oithing bean in an intercorping system; and to determine agronomical planting densities attained by planting 2 seeds per hill spaced at 10 Data from www.nature.com In AGRIS collection since: 2005 A crop of maize variants A transposon in <>tbi-tbi-t/i> drove maize Atmast Suwanaft; Jarong Rungchuang; Sompom Thongdang (Kasetset Univ, Bangkok (Thailand). Faculty of Agriculture. Department of Soil Science) Science) Us processor rejects maize that EU won't trait | Pajic, Z., Maize Research Institute Zemun Pojie, Belgrade - Zemun (Maize Research Institute Zemun Pojie, Belgrade - Zemun (Serbia); T Research Institute Zemun Pojie, Belgrade - Zemun (Serbia); Pavlov, The breeding programme on speciality maize with specific traits was initial material was collected, new methods applying to breeding of a enhance and improve variability of the initial material for breeding the | odorovic, G., Institute for Medicinal Plant Research Dr Josif Pancic, Belgrade (Serbie); Srdic, J., Meize M., Meize Research Institute Zemun Polje, Belgrade - Zemun (Serbia) established at the Maize Research Institute Zemun Polje, Belgrade - Zemun (Serbia) several decades ago. The peciality maize, I.e., popping mazz, sweet maize and white-seeded maize, were introduced. The aim was to | Publication date: 2004 Language: sr Deba provider: Narodna bibliotka Srbije - National Library of Serbia (NLS) Type: TEXT |
| Research Center, Bako (Ethiopia))) The experiment was conducted at Bako research center in 1998 and 1999 cropping seasons. The objectives of the study were to determine agronomically optimum plant • Biotechnology in maize breeding productivity. Factorial combinations of three maize plant populations (50, 75 and 100% of optimum planting densities attained by planting 2 seeds per hill spaced at 10 • Biotechnology in maize breeding In AGRIS collection since: 2005 • A crop of maize variants • A crop of maize variants Effects of groundnut and green manure legumes intercropped to maize on yields of the intercrop maize, weeds and moisture of black clay soil • Us processor rejects maize that EU won't term Amnat Suwanaft; Jarong Rungchuang; Somporn Thongdang (Kasetsart Univ., Bangkok (Thailand). Faculty of Agriculture. Department of Soil Science) • Us processor rejects maize that EU won't term | | | Breeding of maize types with specific traits at Maize Research Institute, Zemun Polje |
| Effects of groundnut and green manure legumes intercropped to maize on yields of the intercrop maize, weeds and moisture of black clay soll domestication domestication domestication y US processor rejects maize that EU won't ta domestication y US processor rejects maize that y US processor rejects maize t | Research Center, Bako (Ethiopia))) The experiment was conducted at Bako research center in 1998 an population of maize and climbing bean in an intercropping system; a productivity. Factorial combinations of three maize plant populations | d 1999 cropping seasons. The objectives of the study were to determine agronomically optimum plant nd to determine appropriate planting time of climbing bean to be grown with maize for high total system | Biotechnology in maize breeding Data from www.nature.com |
| Annat Suwanank, salong Kungoruang, Somport monguang (Kasatsan Onix, Bangkok (Inawano), Pacuky of Agriculture. Department of Soli Science) | | | * domestication |
| | | | |

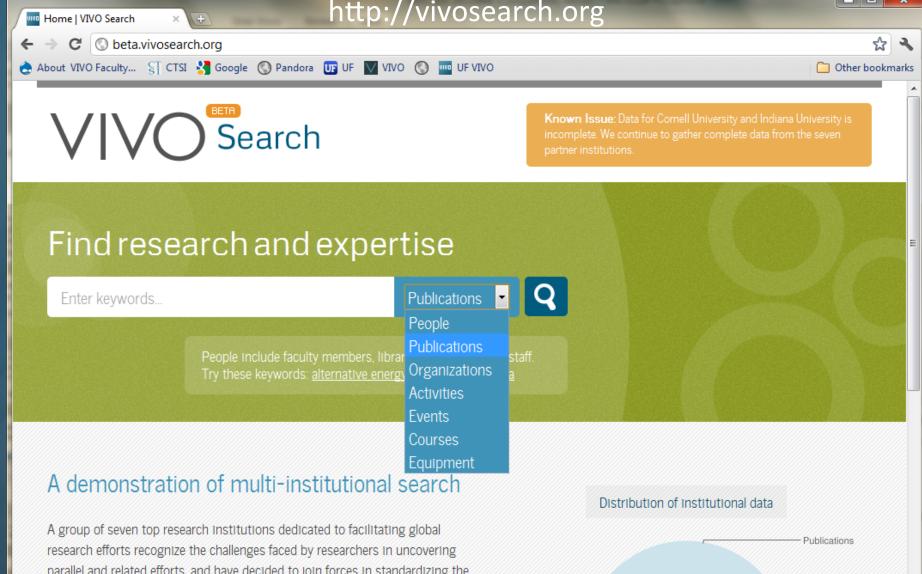
Field experiment was conducted on Takhi soil series (Typic Calciustolis) in a farmer's field in Pakchong, Nakhon Ratchsims, to examine effects of intercoopping groundnut and green manure legumes to maize on the yields of the intercop maize, weed incidence and moisture status of the soil, as an effort to find green-manure legumes that could be intercopped to maize with minimum detrimential effect on the intercop maize and could continue to grow during the fallow period after harvest of the int ...

In AGRIS collection since: 2005

http://agris.fao.org/openagris/searchIndex.do?query=maize

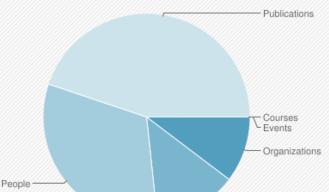
2011 VIVOsearch LOD approach





parallel and related efforts, and have decided to join forces in standardizing the way institutional data gets published. Each institution uses the VIVO software to manage and publish up-to-date information about researchers and their activities.

This website provides a working example of how a multi-institutional search functions, allowing you to search across all seven partner institutions and across all disciplines to find people and information that could dramatically





Known Issue: Data for Cornell University and Indiana University is incomplete. We continue to gather complete data from the seven partner institutions.

Find research and expertise

microbiome

People



People Publications ¹⁴ Organizations ⁶ Activities ³ Events Courses Equipme

13 results

Ley, Ruth E.

... Ley, Ruth E. A core gut **microbiome** in obese and lean twins Ley, Ruth E. Succession of microbial consortia in the developing infant gut **microbiome** Ley ...

Ornell University

Wang, Gary Ping-Chuan

... Wang,Gary Ping-Chuan AST PROF Wang,Gary Ping-Chuan Wang,Gary Ping-Chuan Human Oral **Microbiome** in Periodontal Disease and Hiv Infection Hcv ...

University of Florida

Wolan, Dennis

... **Microbiomes** and Disease Associations between human microbial environments and the onset of human diseases are rapidly materializing. Elucidation ...

🗲 The Scripps Research Institute

Heath,Andrew C

... Spencer T. Olin Prof of Psychology in Psych Heath, Andrew C Heath, Andrew C A core gut **microbiome** in obese and lean twins. Heath, Andrew C Peer ...

🗱 WashU in St. Louis School of Medicine

| Re | esults by Institution | |
|----|--|---|
| UF | University of Florida | 7 |
| ⊜ | Cornell University | 2 |
| \$ | The Scripps Research Institute | 2 |
| 暾 | WashU in St. Louis School of Medicine | 2 |
| Ŧ | Harvard University | |
| ψ | Indiana University | |
| 0 | Ponce School of Medicine | |
| ⊜ | Weill Cornell Medical College | |

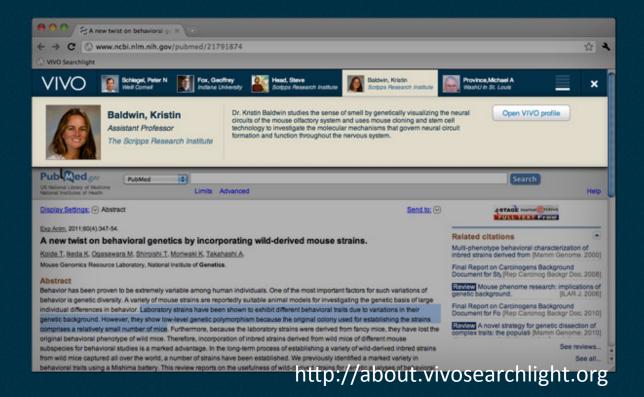
Results by Type

Faculty Member

13

VIVO Searchlight

Searchlight is a small app that automatically shows you VIVO profiles related to the page you're reading.



VIVO



Search for Investigators at Multiple Institutions

| | Text only |
|---------------|---|
| 🗌 Display map | O Text and UMLS concepts |
| | \odot UMLS concepts (including support for boolean search using &, , and !) |

Search

Current Status

- SPARQL endpoints queried: 10
- Institutions indexed: 19
- Total persons indexed: 124,945
- Total publications by those persons indexed as part of their profile: 1,325,716
- Total co-author pairs (two people on the same paper): 2,036,131
- The harvesting times listed below are the times required to interrogate the respective SPARQL endpoints and cache the results locally at lowa.

| Currently Harvested Sites | Platform | Harvesting Time |
|---|----------------------------|--------------------|
| Chicago Women in STEM (Argonne National Lab, Fermilab, Northwestern, U. Chicago) | SciVal Experts and VIVO | 1:11:12 |
| Cornell University | VIVO | 38:24 |
| Duke University | VIVO | 1:07:42 |
| Harvard University | Profiles | 1:22:42 |
| Indiana University | SciVal Experts and VIVO | 17:07 |
| Northwestern University | SciVal Experts and VIVO | 1:43:53 |
| Oregon Health Science University | SciVal Experts | * |
| State University of New York – REACH (Stony Brook University, SUNY College of Optometry, SUNY Downstate Medical Center, University at Buffalo, Upstate Medical University) | VIVO | 0:44 |
| University of California, Davis | SciVal Experts and VIVO | 19:43 |
| University of California, San Francisco | Profiles | * |
| University of Florida | VIVO | 1:19:40 |
| University of Iowa | Loki | 6:01 |
| Note: sites in italics are currently harvested by means other than SPARQL queries on LOD. | | |

VIVO

http://research.icts.uiowa.edu/polyglot/

CTSAsearch

Search for Investigators at Multiple Institutions

- Text only
- Text and UMLS concepts

UMLS concepts (including support for boolean search using &, |, and !)

microbiome

Search

UMLS concepts recognized in your query:

Environment C1956108 – microbiome

Map view of this query

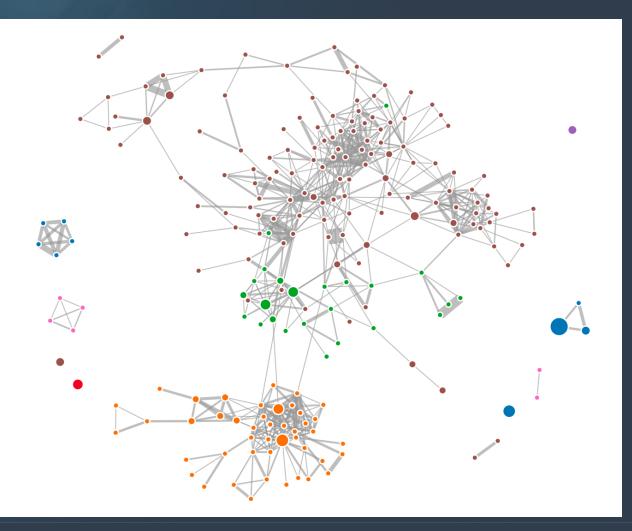
Search Results: microbiome C1956108

Result Count: 296

| Person | Title | Email Address | Phone Number | Score |
|-----------------------|---|-----------------------|----------------|-------|
| Ruth Ley | | rel222@cornell.edu | | 74.39 |
| <u>Maureen Hanson</u> | Liberty Hyde Bailey Professor of Plant Molecular Biology | mrh5@cornell.edu | | 52.24 |
| <u>Volker Mai</u> | Assistant Professor | vmai@epi.ufl.edu | (352) 273–9398 | 48.85 |
| Josef Neu | Professor | neu@ufl.edu | (352) 273–8985 | 46.53 |
| <u>Susan Lynch</u> | Associate Professor in Residence | susan.lynch@ucsf.edu | 415-476-6784 | 45.53 |
| Yvonne Huang | Assistant Professor | yvonne.huang@ucsf.edu | 415-476-9456 | 43.26 |

http://research.icts.uiowa.edu/polyglot/

'Microbiome' network graph



http://research.icts.uiowa.edu/polyglot/

Multi-institutional scenarios

- Multiple campuses of one university
- University and federal lab connections – E.g., Colorado
- Consortia
 - 60 NIH Clinical & Translational Science Awards adopted VIVO as an ontology standard in 2011
- International
 - Instituto Interamericano de Cooperación para la Agricultura (IICA)
 - AgriVIVO.net



International engagement





PROGRAM

UPDATES

Home » About » Announcements »

VIVO joins CASRAI in advancing research interoperability

Posted by Asha Law on Mon, 2012-04-23 09:18

The Leaders of the VIVO Project team (VIVO @) and the Consortia Advancing Standards in Research Administration Information (CASRAI) are today announcing a collaboration to advance a common global approach to research interoperability.

VIVO is an open source ontology and software system designed at Cornell University for researchers and used in many universities in the USA that has attracted interest more widely internationally. It is based on the Semantic Web / Linked Open Data



| Taxonomy category | Description of role |
|---|--|
| Study conception | Ideas; formulation of research question; statement of hypothesis. |
| Methodology | Development or design of methodology; creation of models. |
| Computation | Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms. |
| Formal analysis | Application of statistical, mathematical or other formal techniques to analyse study data. |
| Investigation: performed the experiments | Conducting the research and investigation process, specifically performing the experiments. |
| Investigation: data/evidence collection | Conducting the research and investigation process, specifically data/evidence collection. |
| Resources | Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation or other analysis tools. |
| Data curation | Management activities to annotate (produce metadata) and maintain research data for initial use and later re-use. |
| Writing/manuscript preparation: writing the initial draft | Preparation, creation and/or presentation of the published work, specifically writing the initial draft. |
| Writing/manuscript preparation: critical review, commentary or revision | Preparation, creation and/or presentation of the published work, specifically critical review, commentary or revision. |
| Writing/manuscript preparation: visualization/data presentation | Preparation, creation and/or presentation of the published work, specifically visualization/data presentation. |
| Supervision | Responsibility for supervising research; project orchestration; principal investigator or other lead stakeholder. |
| Project administration | Coordination or management of research activities leading to this publication. |
| Funding acquisition | Acquisition of the financial support for the project leading to this publication. |

CASRAI contributor roles working group

Harvard-Wellcome draft contributor role taxonomy



International engagement



Strategic partnership of euroCRIS and VIVO 23 November 2011 euroCRIS

euroCRIS, a not-for-profit scientific association registered in the Netherlands, and the leaders of the project team of VIVO, an open source Semantic Web sofware application originally developed at Cornell University, have entered into a strategic partnership.

euroCRIS (www.eurocris.org) is furthering the implementation and linking of Current Research Information Systems (CRIS) based on the Common European Research Information Format (CERIF) - commonly indicated with the acronym CERIF-CRIS - and promotes best practice in CRISs, spanning the field from raw experimental and simulated data through research management systems to research publications.

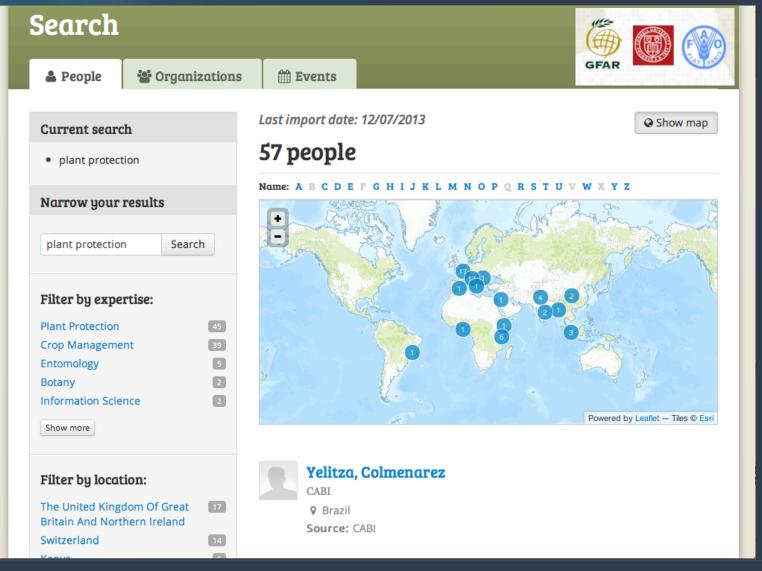


AgriVIVO

| 🐚 AgriVIVO | ном | 1E SEARCH | TOOLS A | ABOUT CONTACT |
|--|---|-------------------|---|--|
| AgriVIVO is a search portal a connections between all act field, bridging across separa and online communities. Vou can search for people, organizatio and events. Read more on how to have included in AgriVIVO. Read our new FAQ, and our terms of use. | ors in the agricult ately hosted direc | tories | Search for periods of the search for period of the search for period of the search of | Search nge", "capacity ment", "information |
| 10 data providers | 788 PEOPLE | 4,74 organizat | | 225 events |

http://agrivivo.net

AgriVIVO



http://agrivivo.net

VIVO

The VIVO-ISF ontology



Ontology definition

 In the context of computer and information sciences, an ontology defines a set of representational primitives with which to model a domain of knowledge or discourse
 – Tom Gruber, 2007



Gruber definition, continued (2)

- The representational primitives are typically classes (or sets), attributes (or properties), and relationships (or relations among class members)
- The definitions of the representational primitives include information about their meaning and constraints on their logically consistent application

Gruber definition (3)

- In practice, the languages of ontologies are closer in expressive power to first-order logic than languages used to model databases
- For this reason, ontologies are said to be at the "semantic" level, whereas database schema are models of data at the "logical" or "physical" level



Gruber definition (4)

Due to their independence from lower level data models, ontologies are used for

- integrating heterogeneous databases
- enabling interoperability among disparate systems, and
- specifying interfaces to independent, knowledge-based services

Otherwise notable

 An ontology is itself embedded as RDF so data becomes self-describing

- Definitions are available via the namespace URI

- The VIVO-ISF ontology reuses significant parts of common ontologies typically organized by domain
- Local extensions roll up into VIVO-ISF for comparison across sites



VIVO ontology goals

- Describe people, organizations, and research resources in the **process** of doing research
- Stay discipline neutral
- Use existing scientific domain terminology to describe **content** of research

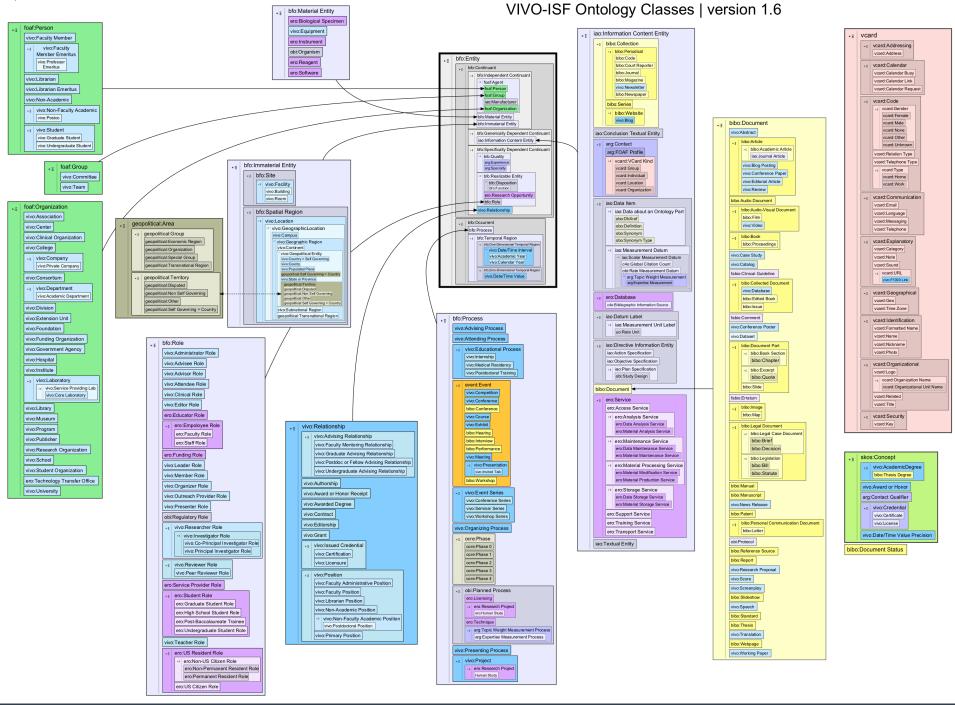


CTSAconnect project (2012-13)

- Aligned and integrated the VIVO 1.5 ontology with the eagle-i ontology for research resources
- The Integrated Semantic Framework, or ISF, includes this work and added some extensions in the clinical domain
- VIVO 1.6+ and eagle-i each use an overlapping set of modules from the ISF
 - Modularity allows selectivity based on local needs

http://ctsaconnect.org, Melissa Haendel, PI

Updated 3/16/14



VIVO-ISF Design Patterns

- Support for information integration

 Vcard ontology
 Authorships
- Support for temporal information
- Support for relationships, roles, and processes
- Connections to geography

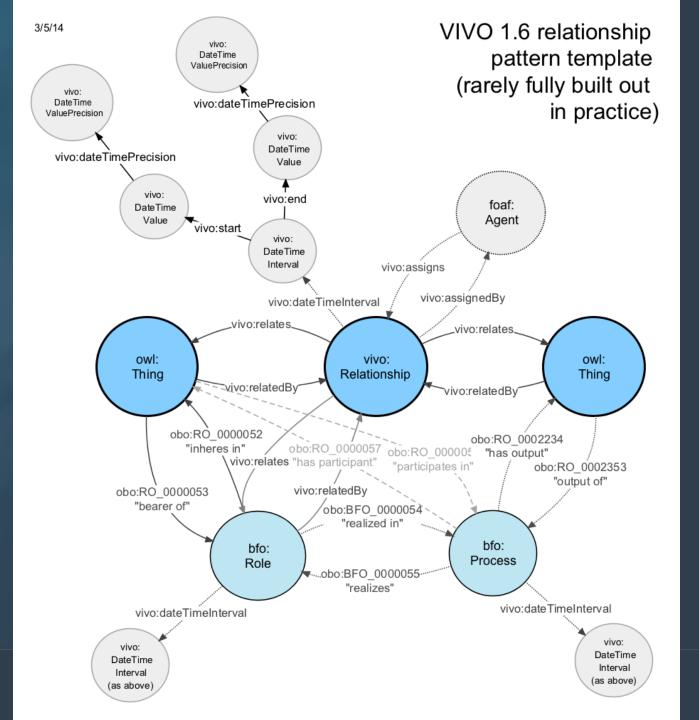


Roles and 'ships

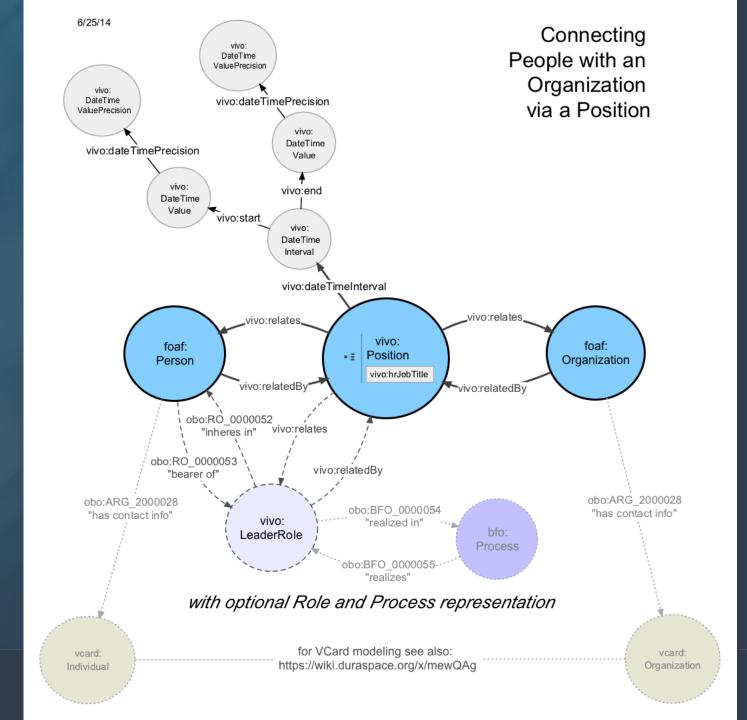
- Typical pattern in RDF is to link two entities via an object property – a simple relationship
- In VIVO-ISF, relationships are entities
- This enables them to hold temporal and/or other information such as author rank
- Often the relationship references a role

Roles and processes

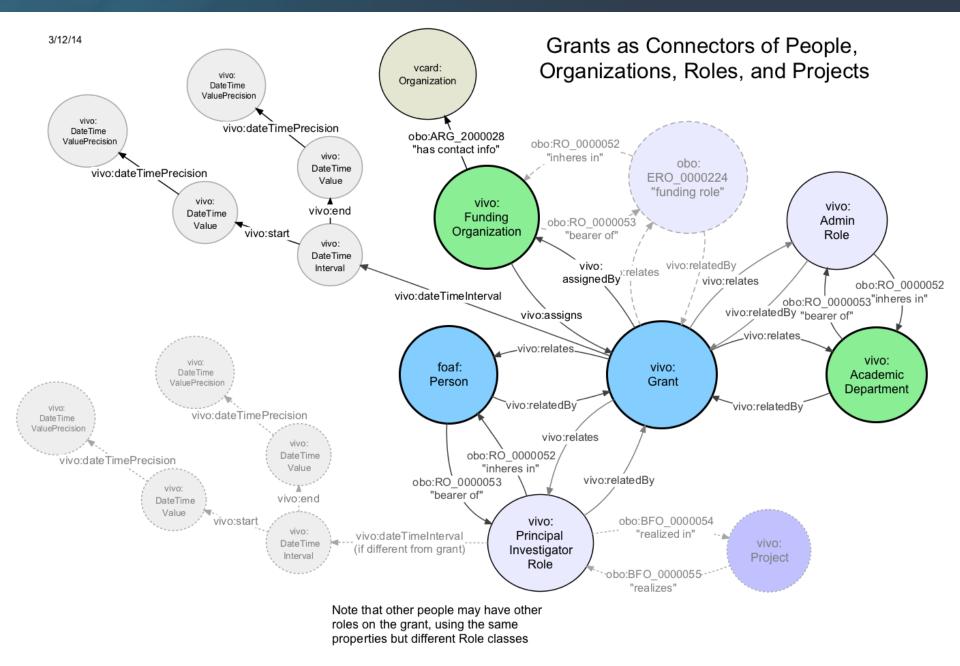
- When you have the information, it's useful to represent the role and the process(es) involved
 - E.g., with a grant, the project and/or investigations
- In VIVO-ISF, the role is distinct (but closely tied to) the relationship
- This model has been applied at Duke for representing humanities scholarship



VIVO



VIVO



VIV

Other ontologies

- VIVO-ISF uses elements from an expanding list of a dozen or more ontologies
- Data interoperability is enhanced by reusing existing ontologies
- The VIVO application adapts to reflect new classes and properties added
 - Some custom forms and "list views" may be needed to support more intuitive editing



Demo: Vitro



What is Vitro?

- The software platform for the VIVO application
 - Approximately 90% of the VIVO code base
- VIVO without the VIVO-ISF ontology and associated customizations and visualizations
 - A blank slate for ontology development and application prototyping

Vitro prototyping exercise

Local Foods

| Photo | Admin Panel Edit this individual Resource URI: http://vitro.vivoweb.org/individual/n5234 Morning Glory produce at Ithaca Farme Opportunity 🔗 | Verbose property display is off Turn on ers Market \mathscr{O} Farmers Market |
|---------------------------------------|---|--|
| Other | | |
| goods distribu | uted 💿 | |
| vegetable green bean sweet corn | | |
| distributor 🕀 | | |
| farm Morning Gl | ory Farm 🖉 🏛 | |
| distribution or | utlet O | |
| farmers marke <u>Ithaca Farm</u> | et n <u>ers Market</u> 🖉 💼 | |

Optional self-study



Optional VIVO & linked data

Learning about VIVO adopters

- Browse any of the <u>publicly available VIVO implementations</u> to compare interfaces, branding, and unique features
- Browse the <u>VIVO Map</u> on our wiki
- Visit vivo.vivoweb.org (ask us for a login)

Multi-institutional search

- Experiment with vivosearch.org
- Try <u>Polyglot</u>, a search across multiple NIH Clinical and Translational Research Awards by Dr. David Eichmann of the University of Iowa

Understanding Linked Open Data (LOD) and basic SPARQL queries

 Exercise: <u>Finding VIVO Data with the University of Florida's public SPARQL</u> <u>endpoint</u>



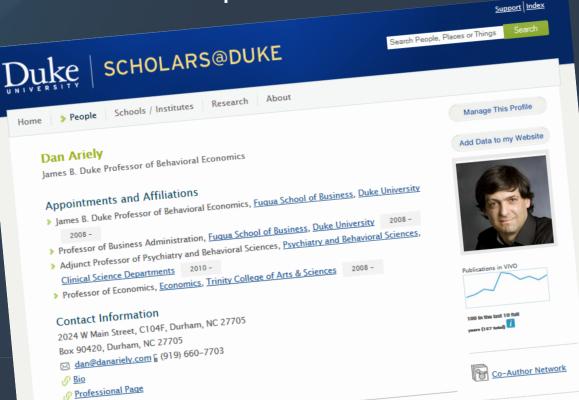
Case Study: Scholars@Duke



VIVO at Duke

- Project team under Provost's Office
- Developers in University IT group
- Elements team in Library
- Currently: 5,685 faculty in 9 schools plus university institutes and centers, 200K+ pubs
- Replacing two legacy systems

scholars.duke.edu

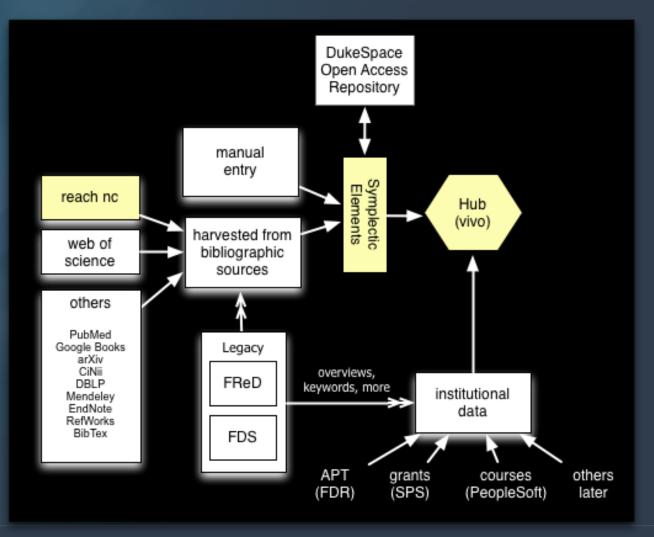


Co-Investigator

Create profiles for all Duke faculty representing their interests and accomplishments

Scholars@Duke: Mission

Data Sources



VIVO

Scholars@Duke publications

Harvest

- Source of articles and keywords
- Identifies authors
- Works well in STEM fields



Manage

- Harvests from REACH NC
- Adds other pubs
- Links to full text publications
- Private profiles

SYMPLECTIC ELEMENTS Display

- Publication list displayed on profile
- Profile data can be repurposed
- Public profiles

Duke | scholars@duke

Rollout Phases



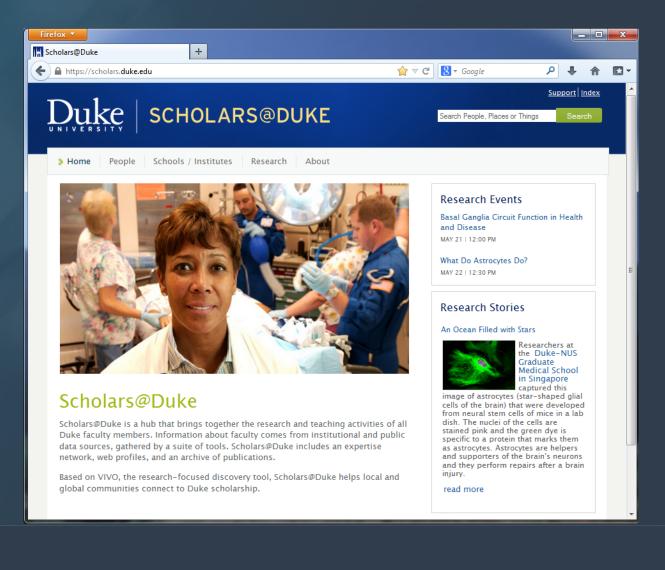
VIVO

2014 Road Map: Phase Two





Scholars@Duke

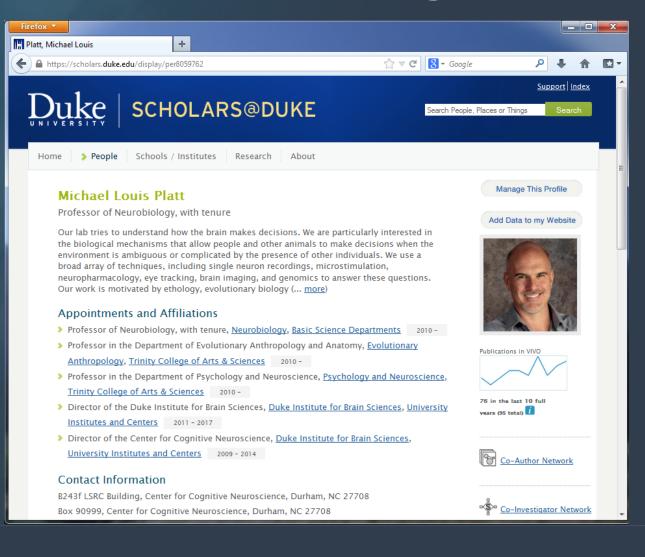


Organizations

| ox 🔻 | | | | | | |
|------------------------------|---|--------------------------------|-----------------------------|-------------|------------|---|
| nools / Institutes | + | a house of | | | | |
| https://scholars-test.oit.du | uke.edu/organizations#org50000299 | ☆ ▼ C | 8 - Google | ٩ | + 1 | |
| | | | | <u>Supp</u> | ort Inde | x |
| Juke 🛛 | SCHOLARS@DUk | (E | Search People, Places or Th | ings | Search | |
| NIVERSITY | | | | | | |
| Home People 🔉 | Schools / Institutes Research Ak | oout | | | | |
| | | | | | | |
| Schools an | d Institutes | | | | | |
| | to see the people and grants. Click the p | lus sign to display the organi | izations | | | |
| within schools. | | | | | | |
| Divinity School | | | | | | |
| Fuqua School of Busin | ess | | | | | |
| Nicholas School of the | Environment | | | | | |
| 👴 Pratt School of Engine | ering | | | | | |
| 🔁 Sanford School of Pub | lic Policy | | | | | |
| School of Law | | | | | | |
| School of Medicine | | | | | | |
| School of Nursing | | | | | | |
| | | | | | | |
| Trinity College of Arts | & Sciences | | | | | |

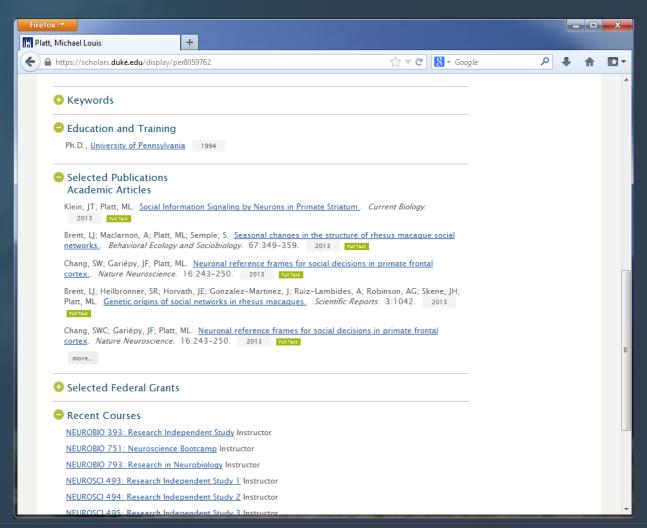


Profile Page

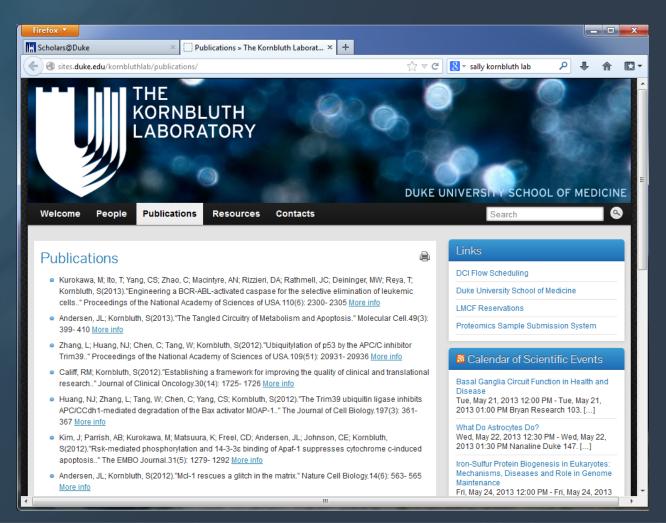


VIVO

Profile Page, part 2

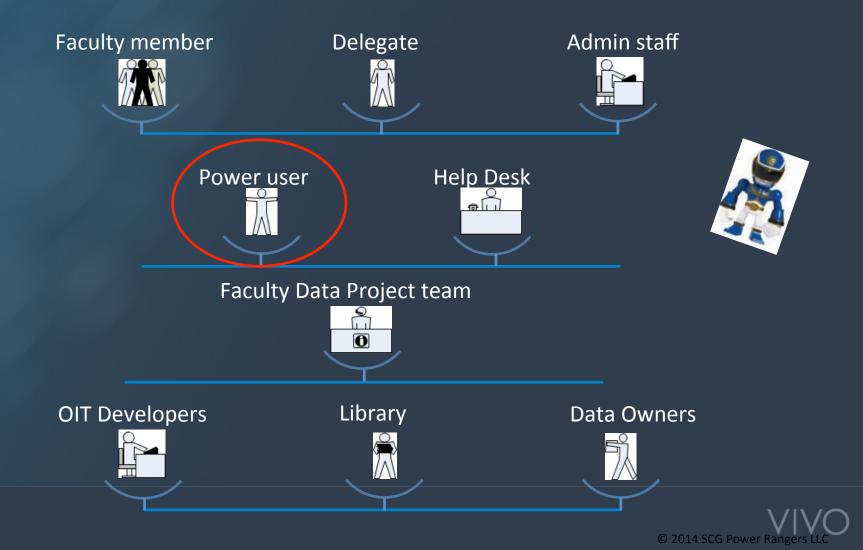


Widget Example



VIVO

Support for Scholars@Duke



Outreach and Training

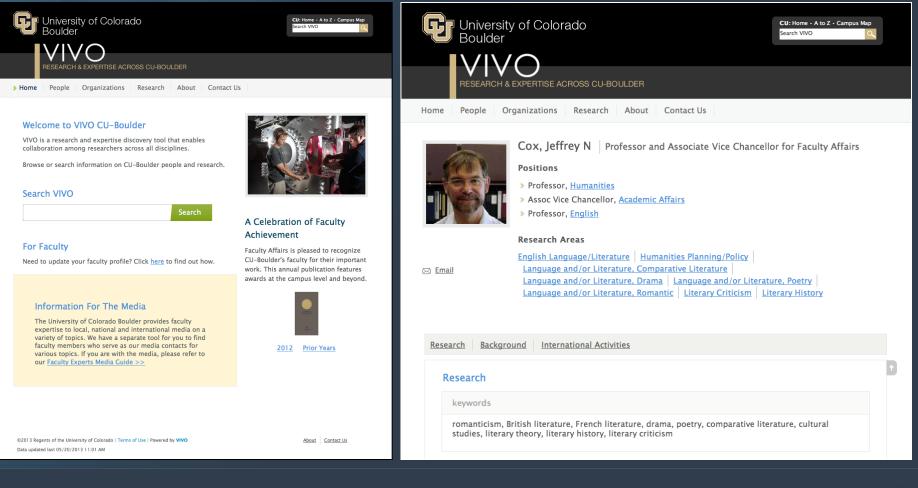
- User Group Meetings
- Power User Training Sessions
- Joint Scholars@Duke & Elements Training Workshops
- Demos at Department Meetings
- Individual Feedback Meetings



VIVO vignettes

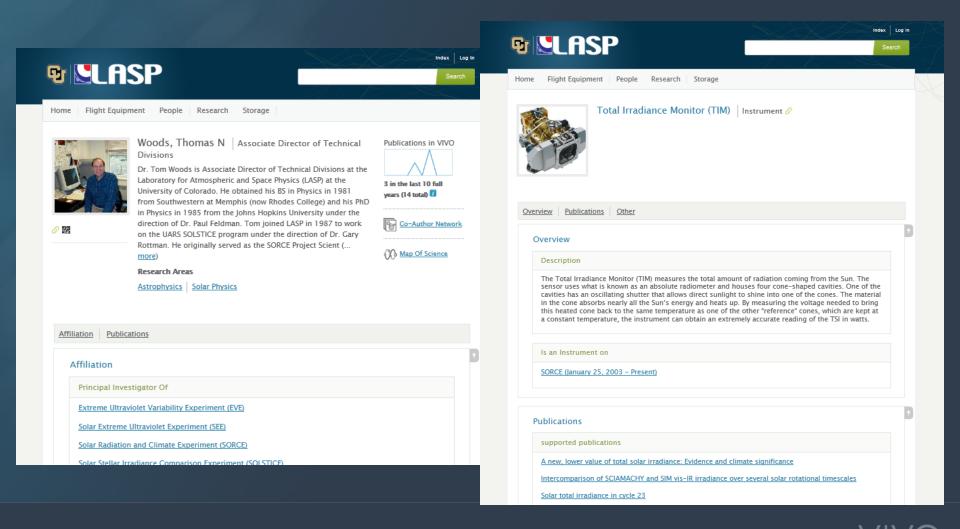


VIVO CU-Boulder



VIVO

Atmospheric & space physics





 \mathbf{x}

Flight Equipment

| Instrument (42) | Space Craft All A B C D E F G H I J K L M N O P Q R S T U V W X Y Z |
|-------------------|--|
| > Space Craft (8) | Cassini Orbiter |
| | POLAR |
| | Solar Dynamics Observatory (SDO) |
| | Solar Mesosphere Explorer (SME) |
| | Solar Radiation and Climate Experiment (SORCE) |
| | Student Nitric Oxide Explorer (SNOE) |

Regional Linked Data Efforts





Leading clean energy innovation



Host institution of 2012 and 2013 VIVO Implementation Fests http://2013vivoimplementationfest.sched.org



CU-Boulder Lessons Learned

- An incremental, value focused approach works for VIVO implementation
- Address faculty concerns as a priority
- Be proactive on data quality
- Build the campus initiative with internal PR
- VIVO's low cost makes it harder to justify resource needs
- VIVO builds campus conversations about Big Data, Linked Open Data, Open Access



vivo.vivoweb.org

- A "VIVO on VIVO" maintained largely by the community
- A resource to find other VIVOs and learn contact names and associated skills



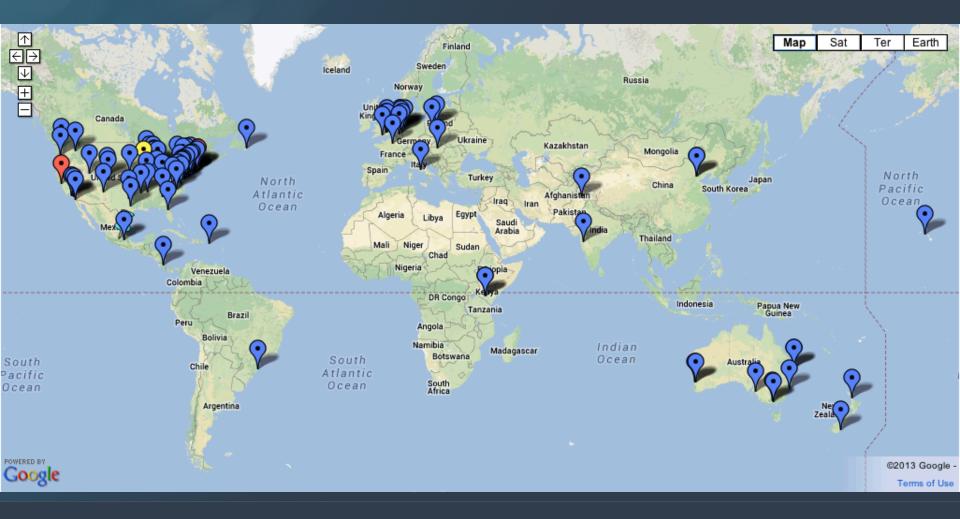
The VIVO Community



VIVO/DuraSpace Partnership

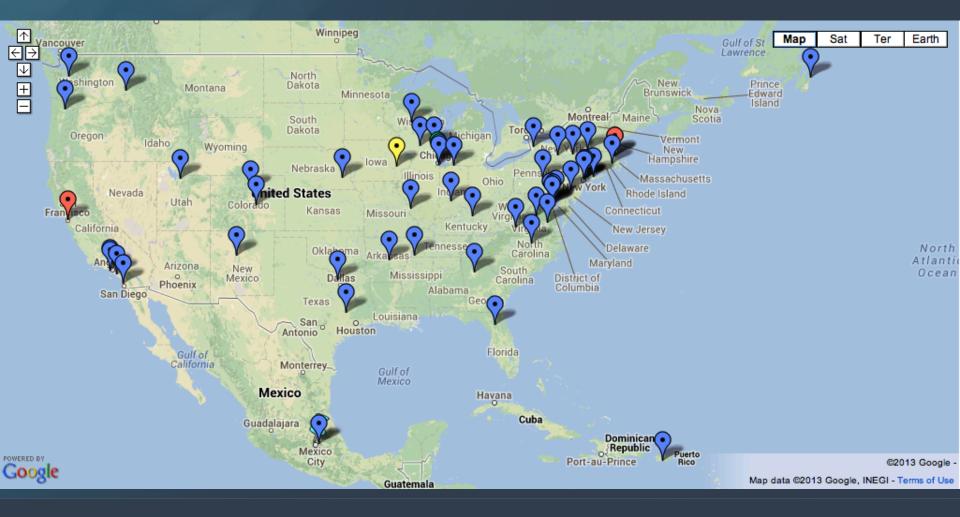
- DuraSpace is a not-for-profit organization supporting the DSpace and Fedora repository platforms and related services
- Proven track record of managing community developed open source projects
- VIVO is 18 months into a 2-year initial startup period
- Several events at the Conference address
 VIVO's relationship to DuraSpace

The VIVO community worldwide



VIVO

VIVO community in North America



VIVO

Where to start?

- Assessing whether VIVO is a good fit for your institution or virtual organization is more about your goals than the technology
- Fundamentally, it's about understanding your needs, VIVO's fit with those needs, and your capacity to sustain the effort



Important indicators

- Do you have institutional sponsors?
 - Starting as a skunk works project is okay but not the best recipe for long-term success
- Does VIVO align with a key institutional initiative?
 - Strategic reinvestment, new academic programs, new senior hires needing information
- Can you marshal resources?

It takes a network

- VIVO is cross-functional
 - Policy, communications, research, library
 - Multiple sources of data
- Requires stakeholder engagement
- VIVO needs to be transparent and fit the research/scholarship culture
 - Not just an "administrative thing"
- It helps to have strong project management
 It's usually obvious whether it's there

Be realistic

- Small, successful pilots targeting one or two constituencies can build momentum
 - Relates closely to CTSA goals but there are equally dynamic initiatives in earth & atmospheric sciences, social sciences, and humanities
- Timelines must allow for ramping up people and technology

Think sustainability

- Loss leader efforts are tempting but if they can't be sustained may backfire
 - E.g., entering a lot of data on behalf of people with no clear update path
- Work with data stewards
 - First, to get access to data you need (public data)
 - To help them better meet your needs via improved APIs or web services
 - To alert them to data issues you may discover
 - VIVO is adept at making problems in source data visible



Reach out

- Interview researchers to learn what they need and want
 - Especially up and coming people building a reputation and more interested in strong online presence
- Create and use an advisory board
- Create a support network

 Duke has "power users"
 Provide materials and training



Benefits across institutions

- Sharing experience provides clarity and new ideas
- Incentives from sharing development, tools, customizations
- Potential data-level connectivity
 - Research is happening increasingly in teams that span institutions
 - Meeting the needs of short and long-term virtual organizations



Use the VIVO community

- We're approachable
- Someone very likely has encountered a similar question or issue before
- Your ideas will be welcome



Use and contribute to the VIVO community resources!

- Wiki
- Listservs
- Regular phone calls
- VIVO events
- Develop local interest groups (e.g., NYC-area sites)

https://wiki.duraspace.org/display/VIVO

VIVO Main Page

Created by Jim Blake, last modified by Jon Corson-Rikert on Aug 03, 2014

New to VIVO?

- VIVO en Español (alpha)
- A short tour of VIVO
- Learning about VIVO
 VIVO map
- Semantic Web resources

VIVO and DuraSpace

- 2013 Incubator project announcement
- Find out about Membership

The VIVO Community

- VIVO adopters
- VIVO working group calls
 - Apps and Tools (bi-weekly on Tuesdays at 1 pm ET, alternating with Ontology)
 Implementation and Development (Thursdays at 1 pm ET every week)
 - Ontology (bi-weekly on Tuesdays at 1 pm ET, alternating with Apps and Tools)
- VIVO listservs
- Calendars and Upcoming Events
- VIVO/SciTS 2014 Conference (August 6 8 in Austin, Texas)
 Other VIVO web sites
 - vivoweb.org and the VIVO blog
 - VIVO project on GitHub
 - VIVO issue tracker
 - VIVO 2013 Conference (presentation downloads)
- VIVO in Social Media

✓ Edit ● Watch I Share Strools ▼

VIVO Releases

- VIVO v1.7 Release Announcement and Release Notes
- VIVO Installation and Upgrade Guides
- VIVO 1.6.2 Release Notes and known issues for VIVO 1.6
- Prior VIVO Releases

Annual Survey

The first annual survey of VIVO implementations is complete; look for a presentation on the results at the upcoming VIVO Conference

Engaging with VIVO

Anyone may read this wiki (and the accompanying $\ensuremath{\mathsf{JIRA}}$ issue tracker)

Anyone may obtain an account on this wiki (and the accompanying JIRA issue tracker)

- To get your DuraSpace ID, send email to sysadmin@duraspace.org, and a real live person will send you a message with a signup link
- To edit issues in JIRA, we just need to connect you with VIVO. Send a note with your DuraSpace ID to request this connection.
- This also means that JIRA issues can be assigned to you!

Collaborations – ORCID

- Open Researcher and Contributor ID

 Attribution for works of any type
- ORCID and VIVO
 - ORCID iD is a property in a VIVO profile
 - VIVO was involved in 3 ORCID Application & Integration grants (Sloan Foundation funds)
 - VIVO 1.7 supports submission and validation of researcher registrations from VIVO, as well as adding VIVO URIs to a researcher's ORCID record

http://orcid.org

For more information

vivoweb.org, vivoweb.org/blog

wiki.duraspace.org/display/VIVO

linkedin.com/groups/VIVO-connect-share-discover

facebook.com/VIVOcollaboration

github.com/vivo-project

@VIVOcollab

Morgan & Claypool publishers

VIVO A Semantic Approach to Scholarly Networking and Discovery

Katy Börner Michael Conlon Jon Corson-Rikert Ying Ding

SYNTHESIS LECTURES ON THE SEMANTIC WEB: THEORY AND TECHNOLOGY James Hendler and Ying Ding, Series Editors

Implementation



Implementation from a technical vantage point

- Options and typical solutions
- Skills and FTE requirements
- Learning about your source data
- Developing an ingest and update strategy
- Leveraging vendor solutions as well as open source communities



Major options

- Physical or virtual hardware
- Choice of OS and base software
- Division of labor
- Approach to data

 Especially for publications
- Staging strategy
- Hosted options?

Physical or virtual?

- Likely depends mostly on your institution's IT environment
 - Physical servers take an up-front investment but may give you more control
 - Virtual servers can usually be scaled according to need
 - Hosted virtual servers can compensate for lack of server administration resources

Choice of OS and software

• Windows or Linux

- Linux more common, but some IT shops have a big Windows investment
- Database MySQL is default, Oracle Enterprise Database an option
- Servlet engine Tomcat is default, Glassfish and others supported
- Web server optional but recommended – Apache HTTP Server

Division of labor

- Skills/roles needed (often from the same person)
 - Sysadmin
 - Database Admin
 - Data conversion/ETL specialist (Java/Python)
 - Data curator
 - Web developer (HTML/CSS)
 - Java developer (optional) for customizing VIVO or adding custom forms
 - User training and support
 - Project management
- Not all need to be full time

Approach to data

- Negotiate with data stewards
- Tools options
 - Harvester updated for VIVO 1.6/1.7
 - Karma, Open (Google) Refine and RDF/semantic tools
 - Python and R, Ruby
- Service providers
- Important to think through data updates, not just a one-time load



Staging strategy

- Allow your techs time to learn Semantic Web concepts and tools
 - Karma is a good way to work with ontologies and RDF and work up to producing VIVO data
 - Enter sample data through the VIVO interface, export it, and study what VIVO itself produces
- Don't start with the most complex data
- Think through what will be interactively updated vs. batch update/replacement
- Test and refine at smaller scales

Resources

- VIVO DuraSpace Wiki
- VIVO Mailing lists
- Weekly dev/implementation and biweekly ontology and apps&tools calls
 - Updates
 - Bug reports and issue discussion
 - Demos of implementations
 - Invited guest presentations
- https://wiki.duraspace.org/display/VIVO

VIVO working groups

- Apps & Tools
- Development
- Engagement
- Implementation
- Ontology

VIVO Implementation Fests

- Four successful events, 2011-2014
- Increasingly about sharing and collaboration more than presentations
- Emphasis on small-group interactions
- Reaching out to related tool providers

VIVO Hackathons

- 2011 hackathon at University of Florida during NIH grant
- March, 2014 event in conjunction with I-Fest
 - Organized by Apps & Tools working group leads Chris Barnes and Ted Lawless
- October 13-15, 2014 at Cornell to grow the committer base for VIVO and tools

4 kinds of open source communities

- Single vendor open source projects
- Development communities
- User communities
- Open source competence centers

What are/will be the salient features of the VIVO community?



Discussion Starting a VIVO in the context of an open source community



Wrap up

Q&A: technical, policy, or strategic issues

