VIVO and the Scholarly Identity Landscape

Internet2 Member Meeting
April 23, 2012
What is VIVO?

• A web-based research discovery tool
  – People plus much more

• Institution-wide, publicly-visible information
  – For external as well as internal audiences
  – A transfer point from Google to official web pages
  – Focus on current activities

• An open, shared platform for connecting communities and campuses
What is a Semantic Web application?

• Provides data readable by machines, not just text for humans
• Provides self-describing data via published (and hopefully shared) ontologies
  – Defined types
  – Defined relationships
• Provides search & query augmented by relationships
• Does simple reasoning
What does VIVO model?

- People, but also organizations, grants, publications, events, facilities, and research resources
  - Most commonly fed from multiple institution systems of record
- Temporally qualified relationships among the above
  - To represent meaningful connections among people and their diverse activities
  - To provide context and navigation from one node of interest to another
- Links to external concepts
Human or machine-readable data at one URI
The VIVO Community
Diversity in adoption

• Small institutions – Ponce, Scripps
• Large research universities – Florida, Cornell, Indiana
• Medical schools – Weill Cornell, Washington U
• Federal government – USDA
• Private sector – American Psychological Association, Wellspring International
• International – UN FAO, Australia, UK, Latin America, China
Value for researchers

• Newcomers building reputations
• Students seeking advisors and mentors
• Student/faculty/senior staff recruitment
• Finding contacts outside your discipline
• Finding facilities, services, and other research resources
Visibility for researchers

- Funding agencies looking for reviewers or contacts for RFIs and workshops
- Reviewers seeking background on investigators
- Corporations looking for expertise, equipment, or advanced materials
- Prospective graduate students and postdocs
- Development and communications staff looking for media commentary
- Administrators seeking input or service
Enter once, use many times

• Provides normalized public data to a range of campus applications
  – Parameterized, filtered queries
  – Search results
• Easily consumable data
  – XML, HTML, JSON
  – Import module for Drupal
  – Open Social gadgets (in progress)
**1221 projects**

» focusing on **295 locations**

» involving **2474 researchers**

» and **1881 organizations**

The College of Agriculture and Life Sciences’ research and outreach address the needs of communities in New York, the United States and around the globe. Explore the many ways in which CALS is making a world of difference.

**Projects by mission focus**

The work of CALS researchers falls into one or more of these project types - all of which are central to the College’s mission.

**Projects by priority area**

CALS has identified four priority areas that are fluid, overlapping, and interdependent, connecting the college’s mission and vision to the areas of Research, Teaching, Extension/Outreach and Development.

**Featured projects**

- Biofuels and Renewable Energy Systems
- Adding Value to the Dairy Industry
- Responding to Global and Regional Climate Change
- Developing Improved Grape Varieties

**International**

671 projects focusing on 158 countries

**26 projects**

- Weather applications for IPM forecasts on farms
- New York Sea Grant Cornell Cooperative Extension facilitates regional sediment management
- Reduced chemical management of golf courses in New York

**Current filters**

- Congressional District
- Year
- International focus
- United States focus
- Priority areas
- Mission focus
- USDA areas
Partnerships – ORCID

“ORCID, Inc. aims to solve the author/contributor name ambiguity problem in scholarly communications by creating a central registry of unique identifiers for individual researchers and an open and transparent linking mechanism between ORCID and other current author ID schemes.”

• VIVO is one way for research institutions to assign ORCIDs
• Exploration of more granular attribution
  – Providing additional incentives for participation
• Joint interest in managing provenance of data from multiple sources
Partnerships – research data

• VIVO/ANDS consortium in Australia
  – Link research data with researcher profiles and publications
  – Harvest with OAI/PMH to national registry

• IMLS grant is extending VIVO with a modular local research data registry tool
  – Extension of previous data staging repository work (Datastar)
Publish Trust Project
American Psychological Association

• Deploying InCommon identities for APA authors
  – Authors authenticate with home institution credentials
• Goal is to bind attributes via InCommon with higher trust, at both APA and home institution
• VIVO is a component in a larger infrastructure for online peer review
VandenBos, Gary R. | Publisher

As the APA Publisher, and the ED of APA's Office of Publications and Databases, Dr. VandenBos is responsible for developing the premier program for the dissemination of psychological knowledge worldwide. His office produces 66 journals, 5 databases and 80 books per year. In addition, the Office of Publications and Databases produces the APA Psychotherapy Video Series and the APA-Style and reference publications which include the Publication Manual of the APA and the Concise Rules of APA (... more)

Research Areas:
- psychotherapy
- schizophrenia

Publications

Selected Publications

CTSAconnect project

• Integrating the VIVO and eagle-i ontologies in a single integrated semantic framework
  – eagle-i is a research resource ontology funded by NIH and developed at Oregon Health & Science University

• Modular ontologies to allow augmenting eagle-i with people and/or VIVO with research resources

• Could potentially run both applications from a single triple store
Linking clinical and research expertise

• Victor Brodsky @ Weill Cornell Medical College inferred expertise from patterns and extent of clinical care
• Stony Brook UMLS web service provides a lookup in VIVO
  – Bridges clinical and research vocabularies
• Part of a larger vision of evidence- and attribution-based expertise
VIVO-euroCRIS cooperation

• euroCRIS is the leading standards effort for research information across the European Union
• Cooperation focuses on interoperation and convergence
• Goal to provide homogeneous access over heterogeneous research information systems
• euroCRIS Linked Open Data Task Group
  – targets a CERIF-VIVO mapping in 2012
  – A collaboration of UN FAO/AGRIS and VOA3R
VIVO-CASRAI Partnership

• CASRAI is a group of leading international research funders and institutions
  – Developing a common data dictionary for data exchange and reuse across the entire life cycle of research
• Partnership will
  – Advance a common global approach to interoperability in the U.S.
  – Advance ontology-centric views of the CASRAI dictionary
• Technology neutral – a diverse network of systems
Macro challenges

• We need dereferenceable identifiers
  – Not just on terminology and people
    • Organizations, journals, places, events
  – Host shared individuals if no authoritative source

• Scaling sameAs assertions and managing duplicates without losing provenance
  – Bilder white paper
Challenges at the micro scale

• We must engage researchers themselves
  – Lattes system in Brazil leave a person’s profile under his or her own control
  – Public sees visual confirmation if information has been verified by an employer or publisher

• Motivations for the individual
  – Do it once
  – Retain control
  – Increase visibility
  – Get credit for additional types of contributions
Thank you.

VIVO Implementation Fest  
May 14-16  
Boulder, Colorado

VIVO 2012 Conference  
August 22-24  
Miami, Florida

Information at  
http://vivoweb.org