Build VIVO in the Cloud

Enabling collaboration and discovery between scientists across all disciplines by providing semantic web-compliant data to the network. See Table of Contents Below. See Slides.

Spotfire
Purpose
2010 VIVO Conference
   Cover Page
   Conference Map
   Program Committee & VIVO History
      Program Committee
      History of the VIVO Project
      Goal of VIVO
   Welcome Letter
   Welcome letter from the Conference & Program Committee Chairs
VIVO Technical Advisory Board & Sponsors
   Technical Advisory Board
   Sponsors
   A Special Thank you...
Schedule of Events
   THURSDAY, AUGUST 12, 2010
   FRIDAY, AUGUST 13, 2010
VIVO Conference Poster
VIVO Workshops
   IMPLEMENTING VIVO AT YOUR INSTITUTION
   VIVO ONTOLOGY
   VIVO DATA ANALYSIS & VISUALIZATION
Scientific Session Information
   THURSDAY, AUGUST 12, 2010
   FRIDAY, AUGUST 13, 2010
About Keynote & Invited Speakers
   Noshir Contractor
VIVO Project Members and Teams

Conference Abstracts

Panels

VIVO: users, interface design, and evaluation.
Implementation of VIVO: experiences from two VIVO installation sites
VIVO Outreach and Adoption: experiences on the local and national level
Next Steps for Research Networking in Science

Contributed Papers

Finding Specialists Using Interface Terminology and Concept-Based Hierarchical Reference Terminology
What faculty members want from a research information system
Configuring and Leveraging a SPARQL End Point for VIVO
CTSAIP: How an Small Project Produced Big Results
Deriving Physicians’ Expertise Profiles Based On ICD9-Coded Encounter Note Logs
Integration Possibilities between VIVO and Google Apps for Education
Standardizing VIVO URLs: How standard is standard?
Drawing Organizational Charts Using VIVO
Digital Vita: Research networking in the context of CV management
From Bench to Bedside and Beyond: Potential Uses of VIVO in an Academic Medical Center Environment
Smarter Campus: Catalyst for Research Collaboration through Optimal Assignment of Resources to Projects
Serving niche and orphan research communities through network extension and augmentation, a user-centered approach
VIVO Development Overview
The Real Who’s Who: User and Computer-based Author Identification
Publications Research and Automated Author Disambiguation: Collaboration between University of Florida and Cornell University
SeerSuite, author disambiguation and VIVO: Building the Semantic Web by crawling
Collexis and VIVO: Paving the Collexis Pathway to the National Research Network
Using VIVO to maximize the h-index

Posters
VIVO Implementation at Indiana University
Implementation and Adoption of VIVO at Washington University School of Medicine
Places and Spaces: Mapping Science
Laboratree –A Web-based Platform for Team Research Collaboration
Using the R Programming Language for VIVO Application Programming
Assessment of Research Impact: A Role for VIVO
“It Takes a Village”: Implementation of VIVO at the University of Florida
Initial comparison of data models in Digital Vita and VIVO
Repurposing VIVO Content with Drupal
Implementation of VIVO at Ponce School of Medicine –the challenges of a small school setting
VIVO Cornell: The Life Cycle of Information
VIVO Users, Interface Design, and Evaluation: Lessons Learned
Development and Application of Subject Knowledge Environment in Chinese Academy of Sciences
Comparative Matrix of Research Networking Tools
Virtual Appliance: releasing without an executable
Creating a collaborative research network for scientists –the example of Mendeley
Drawing Organizational Charts Using VIVO
VIVO Development Road Map Poster
Weill Cornell Medical College VIVO Implementation

Spotfire
For Internet Explorer Users and Those Wanting Full Screen Display Use: Web Player Get Spotfire for iPad App

Media, iframe, embed and object tags are not supported inside of a PDF.

Purpose
Workshop Goals
This 1½ day workshop brings together programmers and users of VIVO specifically and other national researcher networking (NRN) services in general. Demonstration of existing approaches, tools, and techniques as well as discussion of synergies, will provide a point of departure for developing value added services to massive amounts of interlinked (semantic web) scholarly data (my bolding).

The goals of the meeting are:
Present and discuss current research, tools, and services for VIVO/NRN.
Identify synergistic collaboration opportunities and challenges.
Outline a course of activities over the next 5 years.

Given the diverse backgrounds of the participants and the goals of the workshop, we will use the first ½ day for brief self-introductions, followed by three 30 minute overview talks that set the stage for the workshop. The day concludes with a discussion of challenges and opportunities and a hosted dinner. The second full day features brainstorming and discussion sessions in different team sizes and combinations. A particular focus on application development—how VIVO and other NRN can support application development, how application development tools be shared, and what kinds of NRN applications and services are technologically feasible and most beneficial for supporting science.

Brief self-introductions:
My Exhibit Entry
Working with Katy's Borner's Suggested Databases: This suggests it might be useful to try to find open data sources for at least some of the maps so others could reproduce and build on the original results: http://sdb.cns.iu.edu and http://sci2.cns.iu.edu are all about open data and open code.
This Pilot (lots of valuable content in PDF files!)
Prepare for the 2011 VIVO Conference, August 24-26, 2011.

2010 VIVO Conference

Cover Page

VIVO Enabling National Networking of Scientists

FIRST ANNUAL VIVO NATIONAL CONFERENCE
August 12 - 13, 2010, New York City
FOR MORE INFORMATION: HTTP://VIVOWEB.ORG
New York Hall of Science, 47-01 111th Street, Queens, New York 11368
Conference Map

Program Committee & VIVO History

Program Committee

Mike Conlon
Conference Chair
Associate Director, Clinical & Translational Science Institute - University of Florida

Kristi Holmes, Ph.D.
Program Chair
Bioinformaticist, The Bernard Becker Medical Library - Washington University
School of Medicine in St. Louis

Katy Börner, Ph.D.
Victor H. Yngve Professor at SLIS,
Adjunct Professor of Informatics and Statistics; Director, Cyberinfrastructure for Network Science Center - Indiana University

Curtis Cole, M.D.
Chief Information Officer, Associate Professor of Clinical Medicine and Public Health - Weill Cornell Medical College

Paula C. King
History of the VIVO Project

The application which supports VIVO was developed at Cornell University in 2003 and implemented in 2004.

In September 2009, seven institutions received $12.2 million in funding from the National Center for Research Resources of the NIH to enable National Networking with VIVO. This work is funded by the National Institutes of Health, U24 RR029822.

University of Florida is leading the grant with six other participating institutions: Cornell University, Weill Cornell Medical College, Indiana University, Washington University School of Medicine in St. Louis, The Scripps Research Institute, and the Ponce School of Medicine in Puerto Rico.

New institutions are expressing interest in joining the network.
Goal of VIVO

Goal of VIVO:
• Improve science by providing the means for sharing and using current, accurate, and precise information regarding scientists’ interests, activities, and accomplishments.
• Foster team science by providing tools for identifying potential collaborators.
• Improve collaboration by creating tools using this information for enhancing new and existing teams.
• Facilitate team science.

What is VIVO?
• A semantic web application that enables the discovery of research and scholarship across disciplines in an institution.
• Populated with detailed profiles of faculty and researchers; displaying items such as publications, teaching, service, and professional affiliations.
• A powerful search functionality for locating people and information within or across institutions.

Welcome Letter

Welcome to the First Annual VIVO National Conference!

We are excited to be part of this amazing opportunity for people from across the country and around the world to come together in the spirit of promoting scholarly collaboration and research discovery. We have a number of exciting activities planned over the next two days.

We are delighted to welcome Noshir Contractor and Jim Hendler as keynote speakers and very much look forward to their presentations about social networks, the semantic web and their impact on the future of science. We also welcome our invited speakers, an impressive array of noted scientists and professionals, who will speak about an exciting range of topics of interest to the field. We’ve had a great number of wonderful contributions to the program and look forward to thoughtful and engaging dialogue over the next two days. Be sure to join us for the poster session and reception Thursday night in the Viscusi Gallery for refreshments and a relaxed atmosphere for conversation and exploration.

This two-day conference is an excellent opportunity for you to meet with VIVO team members from participating institutions, and offers an open and collaborative environment to share ideas. The conference promises to be
an opportunity for everyone from across the research and scholarly spectrum to discuss topics related to adoption and implementation of VIVO and the opportunities created by the national network of scientists.

Enjoy the conference!

Best regards,

Mike Conlon
2010 VIVO Conference Chair

Kristi Holmes
2010 VIVO Program Chair

VIVO: Enabling National Networking of Scientists is supported by NIH Award U24 RR029822.

Welcome letter from the Conference & Program Committee Chairs

Note: This was hidden from view in the original PDF!

It is our pleasure to welcome you to the First Annual VIVO National Conference in New York on August 12-13, 2010. This year's conference will draw attendance from academic institutions and organizations supporting the development of scientific collaboration from all over the world.

Sessions include presentations by academic scholars and researchers regarding VIVO and collaboration in science. Topics will include adoption and implementation of VIVO and opportunities created by the national network of scientists.

The 2010 VIVO National Conference will deliver up-to-date technical information on VIVO as well as providing you the opportunity to discuss and exchange ideas on the Semantic Web and its impact on science. This conference is for everyone interested in VIVO including scientists, developers, publishers, funding agencies, research officers, collaborators, institutional representatives, and students.

We have a great line-up of speakers including:
• Noshir Contractor, Jane S. & William J. White Professor of Behavioral Sciences in the School of Engineering, School of Communication and the Kellogg School of Management at Northwestern University, USA.
• Jim Hendler, Professor, Computer Science and Cognitive Science, Rensselaer Polytechnic Institute, and Tetherless World Constellation
• And many more!

This two-day conference is an excellent opportunity for you to meet with VIVO team members from all participating institutions and creates an opportunity for networking, collaboration, and sharing of ideas of the possibilities of using VIVO data in rich, meaningful ways. The conference will offer workshops and tutorials: Implementing VIVO at Your Institution, the VIVO Ontology, and VIVO Data Analysis and Visualization Services: How to Program, Extend, and Utilize. Feedback sessions will engage participants in requirements gathering and brainstorming regarding future network services. Contributed papers and panels will allow others to share their work and the poster session reception offers everyone a relaxed atmosphere for discussion and collaboration.

As always, we extend a warm welcome to all our colleagues in the industry who share our interest in improving collaboration in science. We look forward to meeting you at the New York Hall of Science in August at what promises to be a most stimulating and enjoyable event!

Mike Conlon
2010 VIVO Conference Chair

Kristi Holmes
2010 VIVO Program Chair

VIVO Technical Advisory Board & Sponsors

Technical Advisory Board

Chris Bizer  
*Free University of Berlin*

Kei-Hoi Cheung  
*Yale University*

Stefan Decker  
*DERI, Galway, Ireland*

Carole Goble  
*University of Manchester, UK*

James Hendler  
*Rensselaer Polytechnic Institute*

Dean Krafft  
*Cornell University Library*

Carl Lagoze  
*Cornell University Information Science*

Barend Mons  
*Netherlands Bioinformatics Centre (NBIC)*

Abel L. Packer  
*BIREME/OPS/OMS Brazil*

Neil Smalheiser  
*University of Illinois, Chicago*

Steffen Staab  
*University of Koblenz, Germany*

York Sure  
*University of Koblenz, Germany; GESIS Liebniz Institute of Social Science*

Griffin Weber  
*Harvard University*

John Wilbanks  
*Creative Commons*

Sponsors

Science - AAAS  
IMO Interface Technology  
Collexis - an Elsevier business  
Elsevier  
PLOS - Public Library of Science  [http://www.plos.org](http://www.plos.org)  
Thomson Reuters  
UF University of Florida - The Foundation for the Gator Nation  
RefWorks - COS
A Special Thank you...

New York Hall of Science  
University of Florida Department of Conferences  
VIVO Conference Program Committee

Keynote & Invited Speakers:  
Jim Austin  
Katy Börner  
Stefano Bertuzzi  
Noshir Contractor  
Lance DeVine  
Ying Ding  
Jim Hendler  
Barend Mons  
Simon Porter  
Howard Ratner  
Robyn Rebollo  
Griffin Weber  
Stephen Williams

Schedule of Events

**THURSDAY, AUGUST 12, 2010**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:00 AM - 4:00 PM</td>
<td>Registration</td>
<td>Auditorium</td>
</tr>
</tbody>
</table>
| 8:00 AM - 12:00 PM  | VIVO Workshops (Breakfast for workshop attendees provided from 7 - 8 AM in the Auditorium Lobby):  
                      • [Implementing VIVO at Your Institution](http://semanticommunity.info/Build_VIVO_in_the_Cloud)  
                      • [VIVO Ontology](http://semanticommunity.info/Build_VIVO_in_the_Cloud)  
                      • [VIVO Data Analysis & Visualization Services: How to Program, Extend & Utilize](http://semanticommunity.info/Build_VIVO_in_the_Cloud)  
                      | Break Out A  
                      | Break Out B  
                      | Break Out C |
| 12:00 - 1:00 PM     | Lunch (on your own)                                                   |                     |
| 1:00 - 2:30 PM      | Opening Plenary Session - Mike Conlon & Kristi Holmes ([Video](http://semanticommunity.info/Build_VIVO_in_the_Cloud))  
                      Keynote Speaker - Noshir Contractor - Using Web Science to Understand and Enable Research Networks ([Video](http://semanticommunity.info/Build_VIVO_in_the_Cloud))  
<pre><code>                  | Auditorium          |
</code></pre>
<p>| 2:30 - 3:00 PM      | Break (Refreshments provided)                                          | Auditorum Lobby     |</p>
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Venue</th>
</tr>
</thead>
<tbody>
<tr>
<td>3:00 - 4:00 PM</td>
<td>Scientific Sessions &amp; Speakers • Semantic modeling for scientists • Panel Discussion - VIVO: users, interface design, and evaluation • Introduction to ORCID - Open Researcher &amp; Contributor ID (Video)</td>
<td>Break Out A B</td>
</tr>
<tr>
<td>4:00 - 5:00 PM</td>
<td>Scientific Sessions &amp; Speakers • Smarter Campus: Catalyst for Research Collaboration through Optimal Assignment of Resources to Projects • CTSAIP: How a Small Project Produced Big Results; Deriving Physicians' Expertise Profiles Based On ICD9-Coded Encounter Note Logs; Finding Specialists Using Interface Terminology and Concept-Based Hierarchical Reference Terminology • Next Steps for Research Networking in Science (Video)</td>
<td>Break Out A B C</td>
</tr>
<tr>
<td>5:00 - 6:30 PM</td>
<td>Reception and Posters (Hors d’oeuvres, beer and wine)</td>
<td>Viscusi Area</td>
</tr>
</tbody>
</table>

**FRIDAY, AUGUST 13, 2010**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Venue</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:30 - 10:00 AM</td>
<td>Registration (Breakfast for conference attendees provided from 7 AM - 8 AM in the Viscusi Gallery)</td>
<td>Auditorium</td>
</tr>
<tr>
<td>8:00 - 9:00 AM</td>
<td>Keynote Speaker - Jim Hendler - What is the Semantic Web really all about? (Video)</td>
<td>Auditorium</td>
</tr>
<tr>
<td>9:00 - 10:00 AM</td>
<td>Scientific Sessions &amp; Speakers • Panel Discussion - Implementation of VIVO • Shared Health Research Information Network (SHRINE); Research networking in the context of CV Management • Integration Possibilities between VIVO &amp; Google Apps for Education; What Faculty Members Want from a Research Information System; and</td>
<td>Break Out A B C</td>
</tr>
</tbody>
</table>

http://semanticommunity.info/Build_VIVO_in_the_Cloud
Updated: Wed, 23 Sep 2015 06:19:59 GMT
Powered by mindtouch™
<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00 - 10:30 AM</td>
<td>Break (Refreshments provided)</td>
<td>Viscusi Area</td>
</tr>
<tr>
<td>10:30 - 11:30 AM</td>
<td>Scientific Sessions &amp; Speakers</td>
<td>Viscusi Area</td>
</tr>
<tr>
<td></td>
<td>• Panel Discussion - Outreach &amp; Adoption of VIVO</td>
<td>Break Out A</td>
</tr>
<tr>
<td></td>
<td>• VIVO Development Overview</td>
<td>Break Out B</td>
</tr>
<tr>
<td></td>
<td>• CTSciNet and VIVO: A Collaboration Aiming to Enhance Translational Research</td>
<td>Break Out C</td>
</tr>
<tr>
<td></td>
<td>• Building an Australian user community for VIVO: Profiling Research Data for the Australian National Data Service</td>
<td>Auditorium</td>
</tr>
<tr>
<td>11:30 AM - 1:00 PM</td>
<td>Lunch (Food provided)</td>
<td>Viscusi Area</td>
</tr>
<tr>
<td>1:00 - 2:00 PM</td>
<td>Scientific Sessions &amp; Speakers</td>
<td>Viscusi Area</td>
</tr>
<tr>
<td></td>
<td>• Design and development of a modular harvester for data ingest</td>
<td>Break Out A</td>
</tr>
<tr>
<td></td>
<td>• Standardizing VIVO URLs: How Standard is Standard?</td>
<td>Break Out B</td>
</tr>
<tr>
<td></td>
<td>• The Real Who’s Who: User &amp; Computer-based Author Identification; Predicting author h-indexing using characteristics of the co-author network; Author disambiguation &amp; VIVO: Building the Semantic Web by crawling</td>
<td>Break Out C</td>
</tr>
<tr>
<td></td>
<td>• STAR METRICS: An Innovative U.S. Government-University Partnership to Create Data Infrastructure to Couple Scientists and Science Funding with Outcomes</td>
<td>Auditorium</td>
</tr>
<tr>
<td>2:00 - 3:00 PM</td>
<td>Scientific Sessions &amp; Speakers</td>
<td>Viscusi Area</td>
</tr>
<tr>
<td></td>
<td>• Serving Niche and Orphan Research Communities through Network Extension and Augmentation, a User-centered Approach</td>
<td>Break Out A</td>
</tr>
<tr>
<td></td>
<td>• Configuring a SPARQL End Point for VIVO; Drawing Organizational Charts Using VIVO; and Publications Research and</td>
<td>Break Out B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Break Out C</td>
</tr>
</tbody>
</table>

http://semanticommunity.info/Build_VIVO_in_the_Cloud
Updated: Wed, 23 Sep 2015 06:19:59 GMT
Powered by mindtouch™
Automated Author Disambiguation: Collaboration between University of Florida and Cornell University
• Collexis and VIVO: Paving the Collexis Pathway to the National Research Network
• The next step in knowledge evolution; Colonization of brains... (Video)

3:00 - 3:30 PM  Break (Refreshments provided)  Viscusi Area

3:30 - 5:00 PM  Closing Plenary Session  Mike Conlon & Kristi Holmes  Auditorium

VIVO Conference Poster

Note: Trying to extract this information from the PDF to a spreadsheets for Spotfire Analytics.

VIVO Workshops

THURSDAY, AUGUST 12, 2010

IMPLEMENTING VIVO AT YOUR INSTITUTION

Instructors: Valrie Davis, Sara Russell Gonzalez, Christopher Case & Alex Rockwell
Marston Science Library, University of Florida
Location: Break Out A Time: 8 AM - 12 PM
Slides
This workshop will cover the necessary steps to implement VIVO at both small and large institutions. Topics will span necessary team skills, equipment requirements, strategies for acquiring administrative and researcher support, marketing and education, and data management and re-purposing. Participants will install and explore VIVO on their personal laptop utilizing a virtual appliance preloaded with sample data. A thumb-drive will be given at the beginning of the workshop containing the virtual appliance, sample communications to institution administration and brandable marketing materials.

Objectives - Attendees will gain an understanding of:

General Overview
- Who uses VIVO and why
- Linked data and the semantic web
- Required team skills and equipment
- VIVO interface and structure

Ontology
- VIVO ontological structure
- Modifying ontology to meet local needs

Data
- Modeling organizational structure
- Identifying local and external datasets (including people, publications, grants)
- Negotiating data access
- Data ingest process

Support
- Data maintenance and support (redundant, manual inputting, data integrity, refreshing)
- Developing institutional support model
- Marketing and outreach

Personalization
- Personalization of local interface
- Re-use of VIVO data

National and Future VIVO
- Participating in the national VIVO network
- National support resources
- Future VIVO development
VIVO ONTOLOGY

Instructors: Brian Lowe - Cornell University, Stella Mitchell - Cornell University, Jon Corson-Rikert - Cornell University, Ying Ding - Indiana University, Shanshan Chen - Indiana University and Yuyin Sun - Indiana University

Location: Break Out B Time: 8 AM - 12 PM

Slides

Please note: No previous knowledge of ontologies in general and the semantic web are needed in order to participate in the workshop. The workshop will walk the participant through fundamental ontology concepts, the VIVO ontology and its design strategy, the role of the ontology in the VIVO application, and use of the ontology for reporting, visualization, and analysis.

Objectives
Upon completion of the workshop, participants will:
• Understand the VIVO ontology and its purpose and goals
• Be able to use VIVO tools to extend the VIVO ontology for individual institutional needs
• Understand the role of ontologies in sharing VIVO data on the national network
• Understand more about collaborative ontology development on the web

Specifically, each part will cover the following topics:

Part 1
• A brief review of ontologies in general and their basic concepts, a tour of the VIVO Ontology using its ontology editor, review of documentation, and the re-use of existing ontologies.

Part 2
• Design strategies
• More advanced features of OWL and reasoning
• Relationship with other ontologies
• Role of the ontology in the VIVO application
• Making local institutional extensions
• Migrating the ontology and data as the ontology changes

Part 3
• Semantic web, introduction to SPARQL query, path finder, privacy issues in social networking
VIVO DATA ANALYSIS & VISUALIZATION

How to Program, Extend and Utilize

Instructors: Micah Linnemeier, Chintan Tank, Nianli Ma, and Katy Börner
Cyberinfrastructure for Network Science Center, Indiana University
Location: Break Out C Time: 8 AM - 12 PM
Slides and Hands-on Workshop Code Samples

Through the VIVO project, high quality academic data from systems of record becomes available in a common format through Semantic Web technologies. Data that was previously difficult to access and combine becomes available to anyone, creating a unique opportunity for academic and industry stakeholders to utilize this data in conjunction with their own areas of expertise. The most compelling uses of VIVO data might come from 3rd-party developers creating analyses, applications, and services that meet their specific needs.

This hands-on workshop aims to empower participants to understand, access, and utilize VIVO data for administrative, commercial, or research purposes. It starts with a brief overview of techniques and workflows used to analyze and visualize temporal, geospatial, topical, and network datasets at a micro, meso, and macro level. Emphasis is on the design of insightful visualizations. Next, we will present the general VIVO architecture and explain and demonstrate different options to access and work with VIVO data and to use or extend VIVO code drawing on Indiana University’s experience with VIVO service development. All services and applications are documented at a level of detail that makes it easy for others to replicate and extend them.

Last but not least, we will showcase different data analyses and visualizations of VIVO data at the individual, institution, and national level such as:

http://semanticommunity.info/Build_VIVO_in_the_Cloud

Updated: Wed, 23 Sep 2015 06:19:59 GMT
Powered by mindtouch™
• **Individual level.** Statistics and ego-centric scholarly networks on VIVO Profile pages.

• **Institutional level.** Analyses and visualizations of funding intake and publication output for departments and centers accessible via the VIVO Index page. Download of relevant data in tabular and network formats for further analysis using the Network Workbench tool.

• **National level.** Visualization of VIVO installations and their profile holdings together with web page access and general VIVO information requests. Plus, services that use VIVO URIs to access data across different VIVO instances.

The workshop concludes with a general question and answer session.

---

**Scientific Session Information**

**THURSDAY, AUGUST 12, 2010**

1:00 PM - 2:30 PM

Opening Plenary Session - Auditorium (All Attendees)

Mike Conlon & Kristi Holmes

Keynote Speaker: Using Web Science to Understand and Enable Research Networks

*Noshir Contractor* - Short Biography on page 14

2:30 PM - 3:00 PM Break in the Auditorium Lobby

3:00 PM - 4:00 PM
Semantic modeling for scientists - Break Out A  
Ying Ding  - Short Biography on page 17

VIVO: users, interface design, and evaluation - Break Out B  
Ellen Cramer, Leslie McIntosh, Nick Cappadona, and Manolo Bevia  
Abstract on page 24

Introduction to ORCID - Open Researcher & Contributor ID - Auditorium  
Howard Ratner  - Short Biography on page 18

4:00 PM - 5:00 PM

Smarter Campus: Catalyst for Research Collaboration through Optimal Assignment of Resources to Projects - Break Out A  
Carl Osipov, Robin Lougee-Heimer and German Goldszmidt  
Abstract on page 31

CTSAIP: How a Small Project Produced Big Results - Break Out B  
Michael Hazard  Abstract on page 28

Deriving Physicians’ Expertise Profiles Based on ICD9- Coded Encounter Note Logs  
Victor Brodsky, Tru V. Tran, Stephen Kessler II, and Curtis Cole  
Abstract on page 29

Finding Specialists Using Interface Terminology and Concept-based Hierarchical Reference Terminology  
Regis Charlot, Curtis Cole, Adam Cheriff, Andrew Kanter, Fred Masarie, Amy Wang, Alina Oganesova and Frank Naeymi-Rad  
Abstract on page 26

Next Steps for Research Networking in Science - Auditorium  
Katy Börner, Steve Leicht, Titus Schleyer, Griffin Weber, Jim Austin & Mike Conlon - Abstract on page 25

5:00 PM - Reception & Poster Presentations in the Viscusi Gallery

FRIDAY, AUGUST 13, 2010

7:00 AM - 8:00 AM Breakfast provided in the Viscusi Gallery

8:00 AM - 9:00 AM

Keynote Speaker: What is the Semantic Web really all about? - Auditorium  
Jim Hendler  - Short Biography on page 14

9:00 AM - 10:00 AM

Panel Discussion: Implementation of VIVO: experiences from two VIVO installation sites - Break Out A  
Valrie Davis, Caerie Houchins, Sunita Koul, Paula Markes, Leslie McIntosh, Sara Russell Gonzalez and Alex Rockwell  
Abstract on page 25
Shared Health Research Information Network (SHRINE) - Break Out B
Griffin Weber - Short Biography on page 19

Research networking in the context of CV Management
Titus Scheyler
Abstract on page 30

Integration Possibilities between VIVO & Google Apps for Education - Break Out C
John Spadaro & Jack Templin
Abstract on page 29

What Faculty Members Want from a Research Information System
David Marshak & Todd Johnson
Abstract on page 27

From Bench To Bedside and Beyond: Potential Uses of VIVO in an Academic Medical Center Environment
John Kairys, Jack London and Kari Steiner
Abstract on page 31

Mapping Scientific Networks - Auditorium
Katy Börner - Short Biography on page 15

10:00 AM - 10:30 AM Break in the Viscusi Gallery

10:30 AM - 11:30 AM
Panel Discussion: VIVO Outreach and Adoption: experiences on the local & national level
Break Out A
Kristi L. Holmes, Ellen J. Cramer, Valrie I. Davis, Sara E. Henning and Michele R. Tennant
Abstract on page 25

VIVO Development Overview - Break Out B
Jonathan Corson-Rikert, Christopher Barnes, Micah Linnemeier, Stephen Williams & Narayan Raum
Abstract on page 33

CTSciNet and VIVO: A Collaboration Aiming to Enhance Translational Research - Break Out C
Jim Austin - Short Biography on page 15

Building an Australian user community for VIVO: Profiling Research Data for the Australian National Data Service - Auditorium
Simon Porter, Robyn Rebollo and Lance De Vine - Short Biographies on page 16 & 18

11:30 AM - 1:00 PM Lunch in the Viscusi Gallery

1:00 PM - 2:00 PM
Design and development of a modular harvester framework for data ingest - Break Out A
Stephen V. Williams, Christopher Haines, Dale Scheppler, Nicholas Skaggs, Christopher Barnes, Narayan Raum and Yang Li. - Short Biography on page 19

Standardizing VIVO URLs: How Standard is Standard? - Break Out B
Mike Conlon - Abstract on page 30
The Real Who’s Who: User & Computer-based Author Identification - Break Out C
Iris Kisjes and Jessica Kowalski - [Abstract] on page 33

Using VIVO to maximize the h-index
Christopher McCarty, James Jawitz, Allison Hopkins and Alex Goldman - [Abstract] on page 35

Author disambiguation and VIVO: Building the Semantic Web by crawling
Pradeep Teregowada - [Abstract] on page 34

STAR METRICS: An Innovative U.S. Government-University Partnership to Create a Data Infrastructure to Couple Scientists and Science Funding with Outcomes - Auditorium
Stefano Bertuzzi - Short Biography on page 16

2:00 PM - 3:00 PM

Serving Niche and Orphan Research Communities through Network Extension and Augmentation, a User-centered Approach - Break Out A
Jonathan Kuniholm and Joshua Sommer - [Abstract] on page 32

Configuring and Leveraging a SPARQL End Point for VIVO - Break Out B
Christopher Barnes, Narayan Raum, Christopher Haines, Stephen Williams, Nicholas Skaggs, Dale Scheppler and Yang Li - [Abstract] on page 27

Drawing Organizational Charts Using VIVO
Alex Rockwell - [Abstract] on page 30

Publications Research & Automated Author Disambiguation: Collaboration between University of Florida & Cornell University
Christopher Westling and Nicholas Skaggs - [Abstract] on page 34

Collexis & VIVO: Paving the Collexis Pathway to the National Research Network - Break Out C
Michael Warden, Christian Herzog, Brian Athey and Teri Grieb - [Abstract] on page 35

The next step in knowledge evolution; Colonization of brains... - Auditorium
Barend Mons - Short Biography on page 17

3:00 PM - 3:30 PM Break in the Viscusi Gallery

3:30 PM - 5:00 PM

Closing Plenary Session - Auditorium (All Attendees)
Mike Conlon & Kristi Holmes

About Keynote & Invited Speakers

Noshir Contractor
Jane S. & William J. White Professor
Northwestern University
Noshir Contractor is the Jane S. & William J. White Professor of Behavioral Sciences in the McCormick School of Engineering & Applied Science, the School of Communication and the Kellogg School of Management at Northwestern University, USA. He is the Director of the Science of Networks in Communities (SONIC) Research Group at Northwestern University. He is investigating factors that lead to the formation, maintenance, and dissolution of dynamically linked social and knowledge networks in a wide variety of contexts including communities of practice in business, translational science and engineering communities, public health networks and virtual worlds. His research program has been funded continuously for over a decade by major grants from the U.S. National Science Foundation with additional current funding from the U.S. National Institutes of Health (NIH), Air Force Office of Research Support, Army Research Institute, Army Research Laboratory and the MacArthur Foundation.

Professor Contractor has published or presented over 250 research papers dealing with communicating and organizing. His book titled Theories of Communication Networks (co-authored with Professor Peter Monge and published by Oxford University Press) received the 2003 Book of the Year award from the Organizational Communication Division of the National Communication Association. He is the lead developer of C-IKNOW (Cyberinfrastructure for Inquiring Knowledge Networks On the Web), a socio-technical environment to understand and enable networks among communities, as well as Blanche, a software environment to simulate the dynamics of social networks.

Jim Hendler
Tetherless World Professor of Web Science
RPI
James Hendler is the Tetherless World Professor of Computer and Cognitive Science, and the Assistant Dean for Information Technology and Web Science, at Rensselaer. He is also a faculty affiliate of the Experimental Multimedia Performing Arts Center (EMPAC), serves as a Trustee of the international Web Science Trust, and is a visiting Professor at the Institute of Creative Technology at DeMontfort University in Leicester, UK.

Hendler was the recipient of a 1995 Fulbright Foundation Fellowship, is a member of the US Air Force Science Advisory Board, and is a Fellow of the American Association for Artificial Intelligence, the British Computer Society and the IEEE. He is also the former Chief Scientist of the Information Systems Office at the US Defense Advanced Research Projects Agency (DARPA) He is the first computer scientist to serve on the Board of Reviewing Editors for Science. In May of 2010, Hendler was named an “Internet Web Expert” for the US government.

James C. Austin

Editor
Science Careers and PI, CTSciNet

Jim Austin is the Editor of Science Careers, the careers-focused publication of AAAS and Science magazine. In this capacity, he started CTSciNet, a community and information Web site sponsored by the Burroughs Wellcome Fund and aimed at bridging the gap between clinical and laboratory biomedical science and enhancing “translation” in scientific research. Jim was trained as a physicist, earning a PhD from the University of North Carolina at Chapel Hill.

Katy Borner

Professor, Director, Curator
KATY BÖRNER is the Victor H. Yngve Professor of Information Science at the School of Library and Information Science, Adjunct Professor in the School of Informatics and Computing, Adjunct Professor at the Department of Statistics in the College of Arts and Sciences, Core Faculty of Cognitive Science, Research Affiliate of the Biocomplexity Institute, Fellow of the Center for Research on Learning and Technology, Member of the Advanced Visualization Laboratory, and Founding Director of the Cyberinfrastructure for Network Science Center at Indiana University. She is a curator of the Places & Spaces: Mapping Science exhibit, http://scimaps.org.

Her research focuses on the development of data analysis and visualization techniques for information access, understanding, and management. She is particularly interested in the study of the structure and evolution of scientific disciplines; the analysis and visualization of online activity; and the development of cyberinfrastructures for large scale scientific collaboration and computation.

She is the co-editor of the Springer book on 'Visual Interfaces to Digital Libraries' and of a special issue of PNAS 101 (Suppl. 1) on 'Mapping Knowledge Domains' published in April 2004. She also co-edited a special issue on 'Collaborative Information Visualization Environments' in PRESENCE: Teleoperators and Virtual Environments, MIT Press (Feb. 2005), 'Information Visualization Interfaces for Retrieval and Analysis' in the Journal of Digital Libraries (March 2005), and 'Mapping Humanity's Knowledge’ in Environment and Planning B (Sept 2007).


She and her colleagues at the Cyberinfrastructure for Network Science Center serve the
• Scholarly Database of 23 million scholarly records, http://sdb.slis.indiana.edu
• Information Visualization Cyberinfrastructure, http://iv.slis.indiana.edu
• Network Workbench Tool and Community Wiki, http://nwb.slis.indiana.edu
• Science of Science Cyberinfrastructure Portal, http://sci.slis.indiana.edu
• Epidemics Marketplace, http://epic.slis.indiana.edu

She is the recipient of many fellowships and awards, including Outstanding Junior Faculty Award, Pervasive Technology Laboratories Fellowship, SBC Fellow, NSF CAREER Award, and Trustees Teaching Award. She is currently PI or Co-PI in funded research: Collaborative Research: Social Networking Tools to Enable Collaboration in the Tobacco Surveillance, Epidemiology, and Evaluation Network (NSF), Modeling the Structure and Evolution of Scholarly Knowledge (James S. McDonnell Foundation), CAREER: Visualizing Knowledge Domains (NSF), Mapping Indiana’s Intellectual Space (21st Century Grant), Network Workbench: A Large-Scale Network Analysis, Modeling and Visualization Toolkit for Biomedical, Social Science and Physics Research (NSF), Towards a Macroscope for Science Policy Decision Making (NSF), and Creative Metaphors to Stimulate New Approaches to Visualizing, Understanding, and Rethinking Large Repositories of Scholarly Data (NSF).

For more information on her research agenda, teaching and other activities, visit: http://ella.slis.indiana.edu/~katy and http://cns.slis.indiana.edu.

Stefano Bertuzzi
Health Science Policy Analyst
National Institutes of Health, Office of the Director

Dr. Stefano Bertuzzi is responsible for the NIH Return on Investment Program, in the Office of Science Policy, Office of the NIH Director, U.S. Department of Health and Human Services. In this position, Dr. Bertuzzi advises the NIH Director on a wide range of health science policy matters related to the impact of biomedical research on knowledge generation, health, wealth, and national competitiveness.

Dr. Bertuzzi is the NIH lead for the STAR METRICS Project, which under the auspices of the White House Office of Science Technology and Policy aims at developing a novel infrastructure to capture the impact of federal R&D investments.

Dr. Bertuzzi has contributed to reforms of biomedical research systems in foreign countries and has been the NIH lead in the bilateral working group with the European Commission to achieve a funding reciprocity policy between the EC and the NIH.

Dr. Bertuzzi has widely contributed to the NIH revision of the peer review system and to the development of a public access policy to NIH funded publications. He is the recipient of several NIH Director’s award, and other national and international awards.

Dr. Bertuzzi received his Ph.D. in Molecular Biotechnology at the Catholic University of Milan, Italy and after a postdoctoral training in the Laboratory of Molecular Neurobiology at the Salk Institute in San Diego, CA., became a tenured Associate Professor at the Dulbecco Telethon Institute in Milan, Italy. He has authored numerous research publications in neurobiology, published in top scientific journals. His research activities led to the discovery of a novel family of homeobox genes involved in the regulation of neuronal axon guidance in the visual system.

Lance DeVine

Research Support Specialist, High Performance Computing and Research Support
Queensland University of Technology (QUT), Brisbane, Australia

Lance De Vine works with the High Performance Computing and Research Support group at the Queensland University of Technology, Brisbane, Australia. This group provides a diverse range of support services for university researchers.
Lance’s duties have included scientific computation, visualization and application development but over the last 2 years has been heavily focused on data management and information science. Lance is presently working on the development and deployment of VIVO at QUT. Research interests include data mining, machine learning, semantic modeling and information representation and visualization.

Ying Ding
Assistant Professor
School of Library and Information Science, Indiana University

Dr. Ying Ding is an Assistant Professor in School of Library and Information Science, Indiana University. Before she worked as a senior researcher at the University of Innsbruck, Austria and as a researcher at the Division of Mathematics and Computer Science at the Free University of Amsterdam, the Netherlands. She completed her Ph.D. in School of Applied Science, Nanyang Technological University, Singapore.

She has been involved in various European-Union funded projects: research-oriented EU projects (EASAIE, OntoKnowledge, IBROW, SWWS, COG, Htechsight, Esperonto, SEKT, DIP, Triple Space Computing), thematic network (Ontoweb, knowledgeweb), and Accompanied Measurements (Multiple). She is very active in many consultancy projects between University and the companies. She has published 100+ papers in journals, conferences and workshops. She is Program Committee for more than 80 international conferences and workshops. She is co-author of the book “Intelligent Information Integration in B2B Electronic Commerce” published by Kluwer Academic Publishers. She is also co-author of book chapters in the book “Spinning the Semantic Web” published by MIT Press and “Towards the Semantic Web: Ontology-driven Knowledge Management” published by Wiley. Her current interest areas include Webometrics, Semantic Web, citation analysis, information retrieval, knowledge management and application of Web Technology.

Barend Mons, PhD.
Scientific Director
Netherlands Bioinformatics Centre (NBIC)

Barend Mons (born The Hague, The Netherlands in 1957, PhD in 1986 at Leiden University, in The Netherlands) is a molecular biologist who turned to bioinformatics in 2000 after a decade of research on the genetic differentiation of malaria parasites.
Next, he spent 3 years of science management at the Research Directorate of the European Commission and the Netherlands Organisation of Scientific Research. During his years with the European and the Dutch research organizations he built up extensive international research management expertise. In that period he was also co-founder of several international research and development initiatives, such as SHARED, The European Malaria Vaccine Initiative (EMVI) and AMANET, a pan African initiative to prepare vaccine testing sites and the Multilateral Initiative on Malaria (MIM).

Barend was co-founder of several academic spin-off companies, including Kreatech diagnostics and Collexis. In 2005, he co-founded Knewco, Inc. Since 2002 he has been Associate Professor in Biosemantics at the Department of Medical Informatics, Erasmus Medical Centre, University of Rotterdam and (since 2005) at the Department of Human Genetics at the Leiden University Medical Centre, both in The Netherlands.

Currently, Barend is a Scientific Director at the Netherlands Bioinformatics Centre (NBIC). He is the initiator of WikiProfessional and an inventor/patent holder of the Knowlet technology. In 2008 he was one of the driving forces behind the Concept Web Alliance, in close collaboration with (a.o.), Jan Velterop, Mark Musen, Amos Bairoch, and Katy Borner.

Summary of output: Peer reviewed articles: 69; Patents: 3; Spin offs in public sector: 6; Spin off activities (private): 3

Selected General Press: Key biology databases go wiki (Nature news) http://www.nature.com/nature/journal...l/445691a.html; Large-scale community protein annotation – WikiProteins; Inaugural Meeting of the Concept Web Alliance (PloS); Concept Web Alliance Hits Ground Running in Bid to Harness Semantic Web for Life Sciences (Bioinform); Literature mining: Speed reading (Nature News) http://www.nbic.nl/uploads/media/Nat...ed_Reading.pdf

Simon Porter
Information Manager (Research), Strategic Analyst
Reporting & Research Innovation
University of Melbourne, Australia
Simon Porter has worked for the University of Melbourne for the past 10 years. Over this time he has gained significant experience on the structure of the university’s administrative information by working on consecutive Student, Finance, HR, and Research Systems implementations. In 2007, Simon designed and helped implement the university’s ‘Find an Expert’ system, a researcher profiling system based on existing administration information sources. Currently, he is working on an internally-focused metrics-based staff profiling system, as well as the creation of the University’s research data registry. Although Simon’s formal title of Information Manager (Research) does not reflect line management responsibility for research information, it does reflect a central purpose in Simon’s activities – that of giving the University a serious addiction to a core set of data on research to be used for multiple purposes.

Howard Ratner

CTO, Executive Vice-President
Nature Publishing Group

Howard Ratner is Chief Technology Officer, Executive Vice-President, for Nature Publishing Group. Based in New York, Howard is in charge of US administration and has global responsibilities for web development and operations, content services, production and manufacturing, and information technology across all NPG products. Howard’s prior positions include Director, Electronic Publishing & Production for Springer and a member of the production staff at John Wiley & Sons. Howard helped launch the DOI, CrossRef, and CLOCKSS initiatives. Most recently he is co-chair of the ORCID initiative -- a community effort to establish an open, independent registry to resolve researcher name ambiguity. He serves on the boards of the International Association of Scientific, Technical and Medical Publishers (STM), Society for Scholarly Publishing (SSP), and CrossRef. He also leads CrossRef Metadata Services and is a member of the STM Future Lab committee.

Related Sites: [http://www.nature.com](http://www.nature.com) and [http://www.orcid.org](http://www.orcid.org)

Robyn Rebollo

eResearch Senior Specialist (Information Management)
Griffith University, Australia
Robyn has over fifteen years experience in information management, data management, online research, library technologies and vendor relations. Her authority in the information industry has been well established, having contributed to the design and development of information and knowledge solutions for business, education, legal and information brokers in the United States and Asia Pacific. Robyn is widely published in research and information journals and is a noted contributor to key trends in the information management industry. Her current role is eResearch Senior Specialist (Information Management) for the eResearch Services, Research Data Services Unit, at Griffith University in Queensland. She assists both individual researchers, and large research projects, analyzing information management needs, designing solutions and supervising the development of solutions to meet research needs. Robyn was the Project Manager for the Griffith University NCRIS Project, which was responsible for the efficient identification and collection of research data records for publishing to Research Data Australia. She is also a project member of the Griffith University and Queensland University of Technology’s EIF Project, which is using Vivo for showcasing researcher profiles and their related research outputs. She is a member of the Special Libraries Association’s Australia and Strategic Librarian Chapters.


**Griffin Weber, MD, PhD.**

Chief Technology Officer  
Harvard Medical School

Dr. Griffin Weber is the Chief Technology Officer of Harvard Medical School (HMS); Director of the Biomedical Research Informatics Core (BRIC) at Beth Israel Deaconess Medical Center (BIDMC); and Assistant Professor of Medicine at Harvard Medical School in the Division of Interdisciplinary Medicine and Biotechnology, Department of Medicine, Beth Israel Deaconess Medical Center. His research focus is in expertise mining and social network analysis. He invented the Harvard Catalyst Profiles website, which contains research profiles for 20,000 faculty that are linked together through both Passive Networks, which are automatically generated based on information known about investigators, and Active Networks, which the users themselves create by indicating their relationships to other researchers. These networks have numerous applications, ranging from finding individual collaborators and mentors to understanding the dynamics of an entire research community. Dr. Weber is also an investigator on Informatics for Integrating Biology and the Bedside (i2b2), an NIH National Center for Biomedical Computing, for which he developed a web-based open source platform that enables a variety of functions including queries of large clinical repositories, visualization of temporal data,
Stephen Williams
IT Expert, Software Engineer
Clinical and Translational Research Informatics Program
University of Florida

Stephen Williams is a graduate of the University of Florida’s College of Liberal Arts and Sciences Majoring in Computer Science. Originally a un-decided major PreMed student, Stephen’s employment with the UF Computing Help Desk lead to him taking a programming class in addition to his usual Premed/MicroBiology classes. The next semester he completed his premed course work and switched his major to Computer Science.

Work at the UF Computing Help Desk gave Stephen a broad overview of the IT landscape from desktop administration & customer support to software development. He accepted a full time Teams Position with the University of Florida Libraries in February of 2006 where his work focused on software development of applications to support Library administration and faculty endeavors. Following a mini-grant from Valerie Davis and Dr. Sara Russel-Gonzalez to bring Cornell’s VIVO program to UF, Stephen was tasked with installing the application. When Cornell and Florida and 5 other institutions teamed up to apply for an NIH Grant that would expand VIVO into a National Network of Scientists, Stephen was named National Implementation IT Support. In February 2010, Stephen joined the Clinical and Translational Research Informatics Program at the Department of Epidemiology, Biostatistics, and Health Policy in the University of Florida’s College of Medicine moving from IT support for VIVO to software development for VIVO focusing on Interfacing with external applications and Packaging.

Stephen is a professional member of the Association of Computing Machinery.

VIVO Project Members and Teams

Note: See Spreadsheet for table and Spotfire for table and visualization of the graphic.

The VIVO NIH U24RR029822 project aims to create a national infrastructure in support of researcher networking.

From September 2009 – August 2011, more than 100 people from seven US institutions — led by Michael Conlon, University of Florida, with input from an external advisory board — will work in diverse teams to synergistically manage all aspects of this socio-technical enterprise.

The map shows the major groupings of team members into implementation (top) and development (bottom), key team leads (colored frames, see legend), members with VIVO-wide team memberships (many linkages), and exclusively local
team memberships (no linkages). VIVO project members and teams are evolving each day: the map is accurate for May 7, 2010. 2010.05.07

Please send comments and questions to Jeni Coffey <coffeyje@indiana.edu> (design) and Valrie I. Davis <vdavis@uflib.ufl.edu> (data acquisition) and Katy Börner (concept).

For more information, visit http://www.vivoweb.org.

---

Conference Abstracts

August 12 - 13, 2010 New York City

Panels

**VIVO: users, interface design, and evaluation**

Ellen J. Cramer, PhD. 1, Nick Cappadona, MS 2, Manolo Bevia, BS 3, and Leslie Mcintosh, PhD. 4

1 Special Projects Lead, Cornell University, Ithaca, NY

2 User Interface Team Lead, Cornell University, Ithaca, NY

3 Interface Designer, Cornell University, Ithaca, NY and

4 VIVO Evaluation Lead, Washington University in St. Louis, St. Louis, MO

Slides

A significant portion of ongoing and proposed technical innovation related to biomedical research revolves around the goal of facilitating the sharing of data, ancillary information, and resources while enhancing collaboration among researchers across a variety of disciplines. VIVO has been designed to address these needs. To assist in this process and understand how VIVO is achieving these goals, the user interface and evaluation teams are working together to assess and improve VIVO.

Specifically, the teams seek to answer the following questions:
1) How well does the software meet the needs of investigators trying to find appropriate people for collaboration and research?
2) How well does the software meet the needs of institutions for learning about their own activities?
3) How accurate and timely are the data at each institution? Is the accuracy related to the techniques used to implement and support the software?
4) What recommendations can be made for improvement?

As software and web site development has become more sophisticated, the focus and terminology has shifted from “user friendly” to user-centered design (UCD) and incorporated usability principles. Designers and developers define the user with use cases, user scenarios, personas, or mental models; each provides a slightly different lens with which to examine a site or the user's needs. The user experience (UX) comprises usability, user analysis, branding, information architecture, and visual design. It is not just about performance or ease of use, it is about the interface providing a positive or affective experience for the user.

As part of the re-design process, the VIVO user interface team chose to create a mental model. The process of creating a mental model involves identifying task-based audience segments and behavior affinities. In addition to this, the entire VIVO team has been involved with writing user scenarios to help direct the development of features and functionality of the application.

During the panel discussion we will share our mental model diagram, results of the user scenario work, and describe how these along with the user testing results have influenced the development of the application. Examples of the wireframes and visuals created as part of this process will be displayed during the presentation. Additionally, results from the first year-evaluation of VIVO will also be presented.

Implementation of VIVO: experiences from two VIVO installation sites

Valrie Davis 1, Caerie Houchins 2, Sunita Koul 3, Leslie McIntosh 4, Sara Russell Gonzalez 5, Paula Markes 6 and Alex Rockwell 7

1 National VIVO Implementation Lead, University of Florida;
2 IT Director, Washington University in St. Louis;
3 Senior Programmer Analyst, Washington University in St. Louis;
4 VIVO Evaluation Lead, Washington University in St. Louis;
5 VIVO National Implementation Coordinator, University of Florida;
6 VIVO National Ontology, University of Florida;
7 VIVO IT Expert, University of Florida;

Slides

Join the implementation teams from the University of Florida and Washington University School of Medicine for a discussion of the implementation process. Information presented during the panel presentation include a discussion of technical issues, data sources, showcase departments, accomplishments and future implementation plans. Ample time will be allowed during the session for questions and discussion.

VIVO Outreach and Adoption: experiences on the local and national level

Kristi L. Holmes 1, Ellen J. Cramer 2, Valrie I. Davis 3, Sara E. Henning 4, Michele R. Tennant 5
1 VIVO National Outreach Coordinator, Washington University in St. Louis;

2 Special Projects Lead, Cornell University;

3 VIVO National Implementation Coordinator, University of Florida;

4 VIVO National Marketing Coordinator, University of Florida;

5 VIVO Outreach Lead, University of Florida;

Slides

Join members of the National Outreach Team for a discussion of VIVO Outreach activities on the local and national level. Information offered during the panel presentation includes a discussion of local outreach strategies for implementation sites, as well as a presentation of national outreach efforts related to adoption and collaboration, the vivoweb.org website, marketing and branding, education and documentation materials, and work with data content providers. Accomplishments and future outreach plans will be outlined and ample time will be allowed during the session for questions and discussion.

Next Steps for Research Networking in Science

Katy Börner 1, Steve Leicht 2, Titus Schleyer 3, Griffin Weber 4, Jim Austin 5 and Mike Conlon 6

1 Moderator, Indiana University;

2 BiomedExperts, Collexis;

3 DigitalVita, University of Pittsburgh;

4 Harvard Catalyst Profiles, Harvard Medical School;

5 CTSciNet;

6 VIVO, University of Florida;

Slides Video

The desire to facilitate and enhance collaboration across scientific disciplines and institutions spans a number of recent efforts by groups in academia, industry and publishing. This panel discussion will offer representatives from many of these efforts an opportunity to present current and proposed work and possible synergies with other efforts. Upon completion of the short presentations, the floor will open for questions from the audience.

Contributed Papers

Finding Specialists Using Interface Terminology and Concept-Based Hierarchical Reference Terminology

Charlot R 1, Cole CL 2, Cheriff AD 2, Kanter AS 3, Masarie FE 1, Wang AY 1, Oganesova A 1, Naeymi-Rad F 1
Finding a particular person with particular expertise is a common problem in scientific collaboration and in medicine. This article describes how a clinical interface terminology, combined with a hierarchical reference terminology (SNOMED®CT), facilitates both clinicians and patients finding physicians – within a physician organization. This search engine addresses several issues: 1) the ability to capture and codify the clinical intent of a search request, 2) the ability to categorize the target population from broad categories to extremely fine expertise, and 3) finally and most importantly, the ability to relate the codified search string to the target population – ranked by relevance – even when no apparent relationship exists.

The Weill-Cornell Physician Organization has a website designed to find physicians using medical consumer or professional terms or physician specialties and this paper describes an enhancement. Interface terminology is used for identifying problems/diagnoses, procedures or physician specialties. It is by this means that clinical search terms are translated into discrete coded concepts, eventually mapping to SNOMED CT at the right level of specificity. Physician specialties are added to the search both as the specialty itself and as groups of diagnoses and procedures that represent the terminological bounds for the health issues addressed by that physician specialty. On the target side, physicians create and maintain a personal profile which consists of similarly-tagged terminology including their specialty as well as any particular expertise in a given disease/health issue or procedure/test. This produces a range of specificity from broad categorizations such as “Pediatrics” to highly specific entries such as “Roux-en-Y bypass” Figure 1 shows how the user’s search terms are translated into medical concepts and specialties and how these are ultimately rendered into SNOMED®CT codes.

The final piece of the puzzle is the ability to relate the codified search request to the most likely physicians based on their profiles. SNOMED CT has a IS-A hierarchy of concepts, which with some minor tweaking allows us to appropriately expand both the search term and the target profiles to increase sensitivity. It also assists the final ranking of potential physicians based on closeness to the original search term. At a specific point in time, this process can be simplified as a single, very large, interface terminology to physician cross-map. For 1,000 physicians and over 300,000 clinical interface terms, cross products result in 1015 to 1020 combinations. Such magnitude is not well handled by current commercial databases, both for response time and scalability. For practical implementation, we use the IMO Terminology Portal ii hosting this cross-map as an in-memory database, where response time is below one millisecond, and concurrent access does not show significant degradation below 80 concurrent requests. This endeavor matches partially formed clinical requests to physicians with particular expertise that may not be found otherwise. We think this technology can be
What faculty members want from a research information system

David W. Marshak, Ph.D. 1 and Todd R. Johnson, Ph.D. 2

1 Medical School and
2 School of Health Information Sciences University of Texas Health Science Center, Houston, TX USA

New information technology for academia is often developed through by administrators and programmers, without input from the end users, and the resulting systems are unpopular, at best. In an effort to adopt a new research information system that will be easy to use and serve its intended purpose, faculty members at UT Houston will be involved from the outset.

As a first step in this process, four, one-hour focus groups consisting of three to five faculty members were led by the authors, a biomedical researcher and an expert on usability of computer programs. Junior faculty members were selected to participate because they had expressed the greatest need for these systems and because they will be the major users in the future. We discussed: Community of Science, the system we had used until 2008; Research Profiles, a system developed by our sister institution, UT Arlington, and recently adopted by the Gulf Coast Consortia; Digital Vita, a system developed at the University of Pittsburgh, and VIVO.

We generated a list of potential benefits from a research information system, consisting, essentially, of everything we do that is not considered scholarly activity. These include: finding collaborators, recruiting students and fellows, locating specialized equipment and other resources, identifying funding sources, preparing grant applications and producing documents needed for grant proposals or those required by the university such as annual reports, research protocols and curriculum vitae. As we were in the midst of a re-accreditation process, this was mentioned prominently. A common theme in the discussions was the need for social networking capability, such as that planned for VIVO, and its possible benefits for assembling and obtaining grant support for multidisciplinary teams. There was a great deal of interest in the potential for matching investigators with donors or industrial partners. The efficient, automated updates of information from multiple sources in VIVO were also popular.

There was an even longer list of problems encountered with the existing systems, including: databases that are incomplete or outdated, systems that are difficult or time-consuming to use, the many, often conflicting, sources of information, the inaccessibility of our own databases and others and the fact that outputs were not generated in useful forms. It is important to emphasize that information about research, in itself, was not seen as sufficient. There was a great deal of interest in generating tangible outputs such as the NIH biographical sketches generated by Digital Vita. Privacy, confidentiality and maintaining priority in scientific discoveries were also major concerns. There would be a great deal of interest in access to bibliographic databases other than PubMed, particularly among those faculty who worked at the interface of biomedical research and other disciplines.

Configuring and Leveraging a SPARQL End Point for VIVO

Christopher Barnes, Narayan Raum, Christopher Haines, Stephen V. Williams, Nicholas Skaggs, Dale Scheppler and Yang Li

University of Florida, Gainesville, FL USA
The keywords associated with the VIVO project: Semantic Web, Linked Data Standard, OpenSource; all of these terms involve openness and accessibility of data. Utilizing the built-in Jena tools, Joseki and JenaToSesame, we can extend the functionality of a VIVO system to include a SPARQL end point. A SPARQL end point allows other systems to query the semantic data store using the SPARQL query language, as defined by the "SPARQL Protocol for RDF".

There are two real world examples of VIVO SPARQL end points. The University of Florida is using the Joseki plug-in for VIVO to open their Jena data store to world. Cornell University is using Sesame linked to VIVO using the JenaToSesame plug-in to replicate th-in to replicate their data to Sesame's Data Store. There are a variety of systems that use these endpoints to collect data from VIVO. UF's Clinical and Translation Research Informatics Program uses the data fromvivo.uff.edu to populate its personnel pages.

CTSAIP: How an Small Project Produced Big Results

Michael Hazard
University of Rochester Medical Center Rochester, NY USA

Introduction
A critical challenge to CTSA Consortium is to identify research strengths and areas in which technologies and intellectual properties (IPs) align, such that we will be able to create opportunities to translate research into new clinical applications or cutting edge first-in-human studies rapidly. The University of Rochester (UR) Clinical & Translational Science Institute (CTSI) is leading an effort to develop a CTSA IP Directory to facilitate consortium-wide collaboration on sharing the data of available technologies and IPs. This is a joint initiative of the UR CTSI Informatics Key Function, UR CTSI Committee for Industry and Foundation Relations, and the CTSA Public-Private Partnership (PPP) Key Function Committee.

This presentation will review how a few people with no funding were able to launch the CTSAIP website, integrate with other national organizations (public and private). Particular emphasis is placed and a pseudo-agile development process that leveraged the knowledge of participating CTSA Technology Transfer offices.

Development Process
We started the development of the CTSA IP Directory in early 2008 with an assessment on the IP databases or datasets maintained by all CTSA institutions at that moment. A small set of common data fields (title, summary, etc.) were identified as the bridging components to facilitate IP data search across institutions. We implemented the search function through a federated data architecture, with each institution being able to maintain the data in their original structure and format (XML files or Excel spreadsheets) and therefore minimized the workload and complexity. iBridge, a standard schema for IP data, is partially implemented. For the front-end, a Google sitemap is updated when institutions add or remove technologies, so to ensure that the site’s index with Google is up to date.

About CTSA-IP
The purpose of the CTSAIP is to aggregate and market technologies from CTSA institutions as well as those of the National Institutes of Health, with the goal of enhancing research activity and private partnerships across the CTSA consortium. The website includes a text-searchable interface and regular, automatic updating with a standardized template to facilitate broad participation by CTSA consortium members. In addition, we are reaching out to other organizations and private industry to automate the exchange of technology data via standardized, XML-based formats. In May of 2010, CTSAIP and iBridge (a Kaufmann foundation funded group) connected our two systems and began exchanging data.

Impact on CTSA Strategic Goals and Local CTSA Activities
Impact on CTSA Strategic Goals (SGs): (1) directly addressing SG3 (consortium-wide collaboration), SG4 (health of community and nation), and SG5 (T1 translational research); and (2) can also facilitate SG1 (research capacity) and SG2 (training and career development).

Impact on UR CTSI: a concrete example to demonstrate collaboration and integration of key function services, which is a unique strength of UR CTSI.

**Deriving Physicians’ Expertise Profiles Based On ICD9-Coded Encounter Note Logs**

Victor Brodsky, MD, Tru V. Tran, BS, Stephen E. Kessler II, BA, Curtis Cole, MD.

Weill Cornell Medical College, New York, NY, United States.

Slides

**Background**

Expertise is one of the most important search fields when trying to find physicians on the Internet. One of the challenges for systems such as VIVO is how to get physicians to identify their expertise and maintain their profile. We used existing ICD9-codes within the Electronic Medical Record to arrive at such results. The initial focus was on obtaining an appropriately ordered list of ICD9 diagnosis codes to accurately describe the professional profile of a given physician focusing on their genuine expertise rather than the incidental diagnoses associated with their practice.

**Methods**

All ICD9 codes were extracted from the Epic EMR database, maintaining the association to the encounter note, patient, and the physician IDs. To identify valid approaches, a sample physician was chosen at random and all 763 of the associated ICD9 codes were manually marked as either “relevant” (184) or “irrelevant” (579) to this physician’s profession.

Prior to attempting to cluster the ICD9 codes into the two groups, several strategies of scoring the ICD9 codes to rank them based on relevance were identified. Using the number of patients and/or encounters for each ICD9 code as a score was the initial approach. The second score was the inverse of the percentage of the physicians who have ever encountered a given ICD9 code, indicating the likelihood that this code has an association with this particular physician. The third score is derived from placing each ICD9 code from a particular physician’s code list onto the hierarchical ICD9 tree and counting the number of each code’s neighbors while grouping them based on the distance from the given code. Each of the scoring strategies needs to be appropriately normalized and integrated to produce a single relevance score for each ICD9 code per physician. To evaluate whether each additional score improved the sorted list of the physician’s ICD9 codes, Hamming distance to the nearest target list was calculated. For the purposes of the evaluation, a target list was defined as having all 184 of its relevant ICD9 codes (as defined by the control) sequentially listed as the first 184 members.

**Results**

In our preliminary findings, the Hamming distance between the algorithmically produced ICD9 list sorted by the described combined relevance score and the nearest target ICD9 list improved from 198 (when based on sorting by the number of patients seen with a given ICD9 code) to 126 when the combined score was used, indicating a 36%
improvement over the simplest method. Additional improvement was demonstrated by the mean, median, and standard deviation of the relevant ICD9 code rankings improving by 52%, 68%, and 26% respectively. 50% of all encountered diagnoses relevant to the given physician’s practice are thereby found in the top 16.5% of the list of all diagnoses rendered by this physician.

Conclusion

Utilizing the electronic medical records to profile physician activity for use in online physician directories and beyond can yield useful associations. A more developed version of this technology should allow prospective and current patients, other physicians, and medical researchers to identify the most likely relevant specialists based on up to date records from the EMR. Additionally, administrators will gain insight into the strengths and gaps of a given institution or department, aiding in strategic and financial planning.

Acknowledgement

Supported in part by NCRR grant 1UL1RR024153-01.

Integration Possibilities between VIVO and Google Apps for Education

John Spadaro 1 and Jack Templin 2

1 Director, Technical Architecture and Outreach, Brown University and
2 Principal, ThoughtCap (consultant working with Brown)

Slides

VIVO and Google Apps for Education are both making strong inroads on campuses across the country. How are these two platforms complementary? How might they integrate? From both a high-level functional and technical perspective, this talk will survey the exciting possibilities.

Standardizing VIVO URLs: How standard is standard?

Mike Conlon, PhD.

University of Florida Gainesville, FL 32653

Slides

Abstract

One of VIVO’s strengths is providing profile information via linked open data. Once a profile is created, it has an address on the Internet. The Uniform Resource Locator (URL) is an Internet standard for addressing content. VIVO profile URLs can be simple (http://vivo.ufl.edu/mconlon), or complex (http://vivo.ufl.edu/individual/n25562/n25562.rdf). Standards for VIVO URLs have not been determined.
This session will provide an opportunity for conference attendees to participate in an open discussion about the approaches to VIVO URLs. Should there be a standard for all VIVO implementations? How simple should/can VIVO URLs be? How do VIVO URLs correspond to other identifiers people might have and share, such as email addresses, network ids, and employee ids? How do VIVO URLs relate to identity management practices at institutions implementing VIVO? What happens if a person changes their identifier or their institution? Following the session, results of the discussion will be summarized and made available on http://vivoweb.org

**Drawing Organizational Charts Using VIVO**

Alex Rockwell

University of Florida Gainesville, FL

Slides

VIVO provides complete information on organizational structures of institutions. Each organization object in VIVO has parent and child organizations. Starting at any particular organization, one can easily use a simple recursion algorithm to traverse the organizations that report up to the starting point. If the starting point is the institution “root”, one can produce an organizational chart for the entire organization. Using Ruby and some open source extensions, simple software has been developed to draw pictures of organizations. Code, algorithm, commentary and sample output will be presented. All code will be available as open source at http://sourceforge.net

**Digital Vita: Research networking in the context of CV management**

Titus Schleyer, DMD, PhD

Center for Dental Informatics, University of Pittsburgh, Pittsburgh, PA

Slides

Identifying collaborators is an increasingly difficult challenge for scientists in an age of growing collaboration in biomedical research. The goal of this project was to design and implement a system for identifying potential collaborators within the six schools at the Health Sciences Center at the University of Pittsburgh. Using a multi-method approach, we developed Digital Vita (DV), a scientific social networking system based on faculty CVs. DV allows researchers to manage their complete CVs, output CV information in a variety of formats, build their social network manually as well as automatically, and maintain and route multiple versions of NIH biosketches. DV is integrated into a scientist’s workflow, reduces work through automatic data acquisition and network effects, and provides secondary benefits. It may be well positioned to serve as a complement to more traditional approaches for identifying collaborators.

**From Bench to Bedside and Beyond: Potential Uses of VIVO in an Academic Medical Center Environment**


* Thomas Jefferson University, Philadelphia, PA, USA

** University of Delaware, Newark, DE, USA

Slides
The VIVO project describes itself as “enabling collaboration and discovery between scientists across all disciplines”. To maximize VIVO’s potential in the areas of biomedical and translational research, that network must clearly include the ranks of physician-scientists and other clinicians.

Ideally, the VIVO system should facilitate the ability of basic science researchers to identify clinical experts who may become future collaborators and supply needed patients for clinical trials, either locally or across institutions. However, reliance on publications, research activities, and grants may not be sufficient to accurately identify all potential collaborators or pools of eligible patients. Thus, inclusion of current clinical interests, involvement in clinical trials, and reporting of clinical volumes may be able to better inform and facilitate successful interactions. The inclusion of data links from clinical research systems (e.g., clinical trials management systems) could provide automated input on all key personnel involved in translational studies.

Furthermore, the needs of physicians to maintain a complete and continually updated CV, to maintain clinical and Continuing Medical Education (CME) records for Maintenance of Certification (MOC), and to maintain that record across institutions may be leveraged to make VIVO a meaningful, life-long tool for meeting these needs. If utilized in this manner, VIVO may also have the potential to provide a robust, nation-wide tracking tool for clinical, educational, and research activities of graduate medical education trainees, as mandated by the Accreditation Council for Graduate Medical Education.

Finally, there is potential for VIVO to facilitate on-going interaction with institutional and program alumni, and to maximize face-to-face social networking opportunities through inventive uses of its activity recording abilities.

**Smarter Campus: Catalyst for Research Collaboration through Optimal Assignment of Resources to Projects**

Robin Lougee-Heimer, Ph.D., IBM Research, Yorktown Heights, NY, USA; Germán Goldszmidt, Ph.D., Somers, NY, USA; Carl Osipov, M.S., IBM Software Group, Orlando, FL, USA; Vladislav Ponomarev, M.S., Anton Zorin, M.S., Ilya Afanasiev, M.S., IBM Russia, Moscow, Russian Federation.

**Slides**

**Abstract**

The Smarter Campus project is a partnership between IBM and academia, to improve the effectiveness of schools by developing better technology, management, measurement and processes. One of the targets of the Smarter Campus project is the research project lifecycle.

The research project lifecycle in most universities is largely a labor intensive, time-consuming, and sub-optimized process. Professors face challenges finding good funding opportunities, identifying collaborators outside their field of expertise, obtaining required resources, and locating needed student assistants. Students have difficulty finding good project opportunities that match their interests, skills, and availability. Administrators have difficulty managing the process in a way that moves their entire institution toward university-wide goals.

Smarter Campus delivers social networking tools, text analytics, and optimization software in an integrated system to support the research lifecycle so that professors, students, and administrators have more time to focus on results and impact. The approach starts with crawling the web to discover and index unstructured data from research publications, grant awards, student social networking profiles and term papers. This information is then stored along with structured data, such as student transcripts, in a data warehouse. Data analytics is used to analyze the unstructured content to
populate research area taxonomy from the project proposal document. Then IBM optimization technology is used to suggest assignments of resources, such as student assistants, to the research projects. A social networking capability is used to display the student and faculty profiles and areas of interest.

The solution relies on search and discovery in a variety of unstructured and semi-structured content (e.g., faculty web pages, research publications, student social network profiles, transcripts) from multiple sources on the Internet (e.g., VIVO) and behind university firewalls. We show how content analytics of qualitative and textually stated preferences in unstructured data can be used in a quantitative mathematical optimization system to advance individual and organization-wide objectives.

**Serving niche and orphan research communities through network extension and augmentation, a user-centered approach**

Jonathan Kuniholm (MS, MID) 1 and Joshua Sommer 2

1 President, Open Prosthetics Project, Durham, NC and

2 Executive Director, Chordoma Foundation, Durham, NC

Slides

“Orphan” conditions are too rare to attract private investment because the commercial sector is simply not able to or interested in developing treatments for small markets. 1 The investment necessary for development of new drugs and devices is too large in comparison to the size of the market for a rare disease. Perhaps tens of millions of Americans suffer from thousands of these conditions, some affecting as few as a dozen patients. The New York Times cites 10 to 20 million patients and 5,000 conditions. 2 The National Organization for Rare Diseases (NORD), 3 lists a smaller number of conditions. Using the regulatory threshold for patient population of 200,000, the NIH lists more than 6,000 such conditions. 4

The authors have personal experience with two of these conditions. Joshua Sommer was diagnosed in his freshman year of college with Chordoma, a rare bone cancer, becoming one of only hundreds of patients living with the disease in the US. 5 Jonathan Kuniholm is one of fewer than 200 servicemembers to lose an arm in recent conflicts, and while estimates vary, joined somewhere between 50,000 and 100,000 people in America missing at least part of one arm, 6 a condition not listed by any of the catalogs of conditions mentioned above.

The needs of these two conditions differ greatly. Chordoma is a malignant bone cancer, for which therapy is primarily surgery and radiation. Ongoing research involves multiple strategies for drug discovery, including developing disease models to enable laboratory research. Prosthetic arms, on the other hand, are needed by people with arm absence as a result of trauma, cancer, and birth defects. Therapies involve surgical interventions as well as prosthetic ones, and the primary cure is a medical device.

What these two conditions share with each other and with the thousands of other orphan conditions is a need to better connect and involve stakeholder groups to research and development. Research is supported by a variety of government sources, and communication is often poor among patients, the funding agencies, researchers and industry.

The two authors have sought to improve collaboration and communication among the communities they support, but have been hampered by a lack of infrastructure enabling such interactions. Social networking and web publishing tools focus on broadcasting information, connecting users to massive and diverse communities, and often lack the laser focus that a niche community requires. Search fails to deliver the highly tailored results required, and interacting with academic papers, patents, poorly marketed products and small conferences is not ideal.

VIVO offers a great opportunity for us, but must be adapted to our needs. Few researchers, many unaffiliated institutions, multiple granting agencies with diverse announcement protocols, and the need to incorporate outside...
information require a novel and dynamic approach. We propose a series of tools to layer information on the VIVO graph, allowing user-generated tags, reviews, wiki, blog and other content. By extending, focusing, and improving the granularity of community-specific information, these networks can improve the quality and quantity of interactions that they afford. Most importantly, our prototypes can be adapted and extended for similar communities.

1 An orphan disease is defined by US statute as any disease or condition impacting less than 200,000 Americans (Orphan Drug Act, US Pub L No. 97-414 (1984, as amended); 21 USC §360aa et seq.).


3 [http://www.rarediseases.org](http://www.rarediseases.org)


6 Amputee Patient Care Service, Walter Reed Army Medical Center.

**VIVO Development Overview**

**Jon Corson-Rikert, BA 1, Christopher P. Barnes 2, Micah Linnemeier, BS 3, Stephen Williams 2, Narayan Raum 2**

1 Cornell University, Ithaca, NY and

2 University of Florida, Gainesville, FL

3 Indiana University, Bloomington, IN

**Slides**

The VIVO Project has created a development plan charting improvements and new features for the VIVO application spanning the life of the grant, through August of 2011, and laying the groundwork for sustainable open-source community development. This presentation will summarize key elements of new development, focusing on three main areas of general interest:

- **How the VIVO application itself is adapting and improving** We present a brief overview of the current VIVO application, including the ability to add local data extensions and address navigation and display customization. Plans for the coming year include leveraging the semantic relationships inherent in VIVO to improve browse navigation, searching, display filtering, and to include visualizations at the individual, site, and national level.

- **How data are acquired, integrated and then shared from a single VIVO for assimilation into national or international aggregations of RDF data** The University of Florida is developing a series of data harvesting and integration tools encompassing both national and local data sources. We discuss the challenges inherent in data integration and how to maximize the return on investment by sharing VIVO Data as widely as possible on other websites within an institution and exposing it as linked data for national or international networking.

- **How the groundwork is being prepared for ongoing software development and data sharing collaborations within and beyond the seven initial partners of the VIVO Consortium**

To be sustainable, the VIVO software needs to become part of a broader community-sourced development, documentation, and extension effort bringing together people from many participating institutions and providing a

[http://semanticommunity.info/Build_VIVO_in_the_Cloud](http://semanticommunity.info/Build_VIVO_in_the_Cloud)

Updated: Wed, 23 Sep 2015 06:19:59 GMT

Powered by [mindtouch](http://mindtouch.us)
focal point for communication and sharing. The http://vivoweb.org site and http://vivo.sourceforge.net provide a coordinated blend of user, implementation, and development information. Plans will be presented for extending Vivoweb.org and the SourceForge site to support an open source software development community.

The Real Who’s Who: User and Computer-based Author Identification

Iris Kisjes 1 and

Jessica Kowalski 2

1 Elsevier, Netherlands and
2 Elsevier, US

Slides

The tasks of disambiguating and identifying scholarly and scientific authors are persistent and pervasive problems across the publishing and academic communities. Scopus author profiles are among a varied group of author identification solutions on the market using multiple approaches to a solution. Both User-based initiatives (which aim to leverage human intelligence) and computer-based solutions (which leverage powerful matching algorithms) have their limitations and advantages. This presentation will compare these two approaches; share the Scopus product team's experience and recent innovations with each.

Publications Research and Automated Author Disambiguation: Collaboration between University of Florida and Cornell University

Christopher Westling, AS 1;

Nicholas Skaggs, BA 2;

1 VIVO Publications, Cornell University, Ithaca, NY and
2 Software, Engineer, University of Florida, Gainesville, FL

Slides

VIVO is uniquely poised to enable the collection and categorization of data ingested from disparate and dissimilar sources. The automated collection of publications data, citation references and other information about scholarly works presents specific challenges for VIVO implementers. When using a methodical approach to search and discover publications for a given institution, we must determine how best to attribute each publication with the correct author. This is of particular importance when two separate citations contain names that are indistinguishable from one another. The ambiguity of author names becomes especially difficult when acquiring and ingesting citation data from several different sources. A human reviewer can often quickly disambiguate two authors with identical name attributes simply by making inferences about associated metadata, but automation is often incapable of making the same decisions. The creation of automated processes to make these assertions is very difficult, and a host of research exists surrounding this problem.

The overall focus of this poster is to represent the complexity of the "disambiguation problem" as it relates to VIVO, document the efforts collectively made to date by member institutions, and discuss directions and expectations for the future. In addition, the cross-site collaboration between Cornell University and the University of Florida will be highlighted in an effort to encourage similar cooperation within the VIVO community as a whole.
Specifically, this poster will document the research methods used to determine the scope and depth of these disambiguation challenges, and will showcase some of the processes created to systematically reduce the number of references that require human intervention for disambiguation. An overview of parallel developments at Cornell University and University of Florida will demonstrate the benefits of collaboration between NIH VIVO member institutions. A brief explanation of the "disambiguation problem" will be presented in basic terms with real-world examples. Descriptions of VIVO's unique features, such as an extensible ontology that provides for the mapping of publications identifiers from different sources, will be explained in basic terms. Future plans to leverage efforts like the ORCID initiative will also be outlined. Graphical results from an early experimental effort to retrieve and match PubMed references with VIVO authors at Cornell will appear alongside similar results from the Harvester product developed at the University of Florida, showing the advantage of carefully constructed processes that enhance VIVO's ability to perform these tasks. The conclusion will summarize the current status of disambiguation efforts, outline plans and expectations going forward, and solicit ideas for potential collaboration with the greater VIVO community.

Mapping the Research Landscape at the Queensland University of Technology

**SeerSuite, author disambiguation and VIVO: Building the Semantic Web by crawling**

Pradeep Teregowda

The Pennsylvania State University

Slides

The growth of the internet presents a challenge in finding experts and authors across various academic communities. SeerSuite was built with the aim of effective dissemination of scientific and academic literature. CiteSeerX, an instance of SeerSuite, hosts a collection of over 1.6 million documents crawled from the web, authored by over 1.5 million authors. However, author disambiguation is an issue and SeerSuite offers various methods for disambiguating author records. This large collection with accurate disambiguation can help identify topic experts and author interests. Utilizing VIVO with author records enables sharing of rich author metadata in SeerSuite for building Semantic Web services.

**Collexis and VIVO: Paving the Collexis Pathway to the National Research Network**

Authors: Collexis: Michael Warden, M.B.A., Knowledge Solutions Leader, Collexis /Elsevier Inc., Ypsilanti, MI, United States; Christian Herzog, M.D., Collexis / Elsevier Inc., Cologne, Germany; University of Michigan: Brian Athey, Ph.D., Chair Designate, Dept of Biomedical Informatics, Professor, Dept of Psychiatry and Dept of Internal Medicine, Director, Academic Informatics, Associate Director, Michigan Institute for Clinical Health Research (MICHR), Ann Arbor, MI, United States; Teri Grieb, Ph.D.,Sr. Director of Research, Managing Director, MICHR, Ann Arbor, MI, United States

Slides

Abstract

Representatives from the University of Michigan and Collexis will demonstrate how working with Collexis' Research Profiles has helped to enable research networking at the University of Michigan, and how Michigan will be using Collexis to connect to VIVO to extend the internal research network into the VIVO National Research Network (NRN).

Building on publication data, grant data and more, Collexis' Research Profiles currently provides innovative ways to address the pressing needs of the academic research environment and has been successfully implemented at the Medical School, and is expanding into other academic disciplines throughout the University. From identifying and evaluating collaborators, matching researchers with new funding opportunities, maximizing efficiency of finding information to strategic research planning, Michigan is using Collexis to be a more dynamic, responsive organization to the needs of the Medical School, collaborating schools and colleges and their faculties. Given the additional focus on
interdisciplinary research and translational research initiatives, Research Profiles enables researchers from different disciplines to find potential collaborators quickly and efficiently.

The University is now launching a connected implementation of VIVO with the assistance of Collexis. This VIVO appliance is automatically built from feeds within Collexis data and provides a method of launching VIVO that requires less employee effort and overhead than the normal implementation. Through this entrance into the NRN, Michigan will begin to extend the connections from researchers inside Michigan out to the rest of the national network while also connecting resources and initiatives located at Michigan to the rest of the scientific community.

Michigan has decided to implement VIVO using the Collexis Profiles solution as an integral part of the institution’s NRN strategy in two ways:

- Utilizing the complete, continuously disambiguated and updated data feed to publish Michigan’s institutional and researcher profiles to the VIVO system.
- Using the Research Profiles application as the front end for Michigan-specific and VIVO-related functionalities (as they become available) to have one interface in front of the user which allows him to connect and network internally but also through the VIVO system with other organizations.

We are beginning to explore the possibility to link into other resource databases such as NIH National Centers for Biomedical Computing (NCBC) UM NCIBI and CTSA resource inventory efforts. Michigan is using the Collexis approach to link researchers at Michigan with potential local and national collaborators and research resources.

Using VIVO to maximize the h-index

Christopher McCarty, James W. Jawitz, Allison Hopkins and Alex Goldman

University of Florida

Gainesville, FL USA

Slides

The h-index is a citation metric that measures scientific productivity and author impact. For a given author, the h-index is the number of articles that have at least that number of citations. Thus an h-index of 10 means an author has ten articles with at least ten citations. A high h-index indicates wide scientific impact. We present the results of an empirical study of the h-index as a function of co-authorship on the Web of Science. We randomly selected 250 authors, calculated their h-index, and the h-index and publication patterns of all their co-authors involved in their h-index articles. We examine the extent to which the structural properties of the co-author networks contribute to the variability in the h-index. Based on these findings we will explore how VIVO can be used to select collaborators that will maximize the h-index.

Posters

VIVO Implementation at Indiana University

Brian Keese, MS; Jon W. Dunn, BA; and Robert H. McDonald, MLIS Library Technologies and Digital Libraries, Indiana University, Bloomington, Indiana, USA

Slides

Indiana University is one of the seven initial implementation sites funded as part of the VIVO project. This poster will summarize work to date, led by the Indiana University Libraries and Digital Library Program, on the implementation of a
pilot VIVO instance to serve as a prototype for the potential implementation of VIVO to support all eight campuses of Indiana University (IU).

The IU implementation team has created two primary instances of VIVO:

**VIVO-on-VIVO.** The VIVO-on-VIVO instance contains profiles for all members of IU’s three VIVO teams: Implementation, Ontology, and Social Networking. Data have been manually entered and maintained, including team organizational information, member positions, research focuses and activities, grants, publications, contact information and educational degrees. Photos have also been uploaded for all team members.

**Development.** The VIVO development instance is in place for the development and testing of automatic data ingest processes. The primary data source being investigated is the Indiana University Information Environment (IUIE). The IUIE is a centrally maintained, enterprise-wide, web-based decision support application and reporting environment. It contains institutional data from a variety of categories, including financial, human resource, and student data, populated from a variety of university information systems, including PeopleSoft HR and Student Information System modules. The data ingested to date include organizational structure; faculty position, title, educational attainments, recent courses taught and awarded grants. Other data are available, including faculty teaching, research, creative, and service activities, and staff and student data. Other data sources under investigation are IU’s institutional repository IUScholarworks, PubMed, and Scopus.

The team’s focus has been on ingesting data from available relational database sources. The main tool in use is the D2R MAP declarative language that is used to describe mappings between relational database schemata and OWL/RDFS ontologies. The mappings can be used by a D2R processor to export data from a relational database into RDF. We are using the prototype processor made available by the Free University of Berlin at [http://www4.wiwiss.fu-berlin.de/bize...map/D2Rmap.htm](http://www4.wiwiss.fu-berlin.de/bize...map/D2Rmap.htm).

Along the way, the team has confronted three major categories of obstacles:

**Cleanliness of available data.** The information within IUIE has inconsistencies due to several factors, such as inconsistent use of data elements across university organizations, and importation of data from multiple legacy systems. These inconsistencies will require manual correction before data can be published.

**Completeness of available data.** Some systems contributing to the IUIE database have only been put into use only recently and as such contain only recent and not complete data. Efforts will need to be made to obtain older data through other means.

**Buy-in from IU community members.** Under institutional data policies, before certain data desired for inclusion in VIVO can be made public, interested parties must give permission, at the data manager, academic school/department, and/or individual faculty/staff/student level. Efforts will need to be made to encourage participation. To this end, the process must be as simple as possible for end users. Automatic data ingest to a large extent will satisfy this requirement, and the more clean and complete the data are the lower the barrier to participation.
Implementation and Adoption of VIVO at Washington University School of Medicine

Kristi L. Holmes, PhD, Sunita B. Koul, Leslie D. McIntosh, PhD, Caerie Houchins, George Joseph, and Rakesh Nagarajan, MD, PhD

Washington University School of Medicine, St. Louis, MO

Slides

Washington University School of Medicine (WUSM) in St. Louis, Missouri, enjoys a rich, nationally recognized tradition of patient care, education, and research. Washington University School of Medicine is one of seven partner institutions that comprise the VIVO Collaboration, an NIH-funded effort to enhance the national networking of scientists. The two WUSM organizations tasked with implementation and support of VIVO locally are Bernard Becker Medical Library and the Center for Biomedical Informatics. These two entities have a history of working together to support research efforts on campus and thus, this project represented a ideal opportunity to leverage skills expertise necessary to implement and support the platform on campus. Two important components of the WUSM VIVO effort are implementation and outreach. Implementation is collecting, storing, and displaying data from WUSM in VIVO. Outreach consists of working with individual departments and research centers on campus and speaking at events sponsored by the Office of Faculty Affairs, the Clinical Research Training Center, and more. Current efforts at WUSM include the addition of complete profile information for Department of Genetics faculty members as well as the integration of demographic, publication, grants, and other relevant information into the local VIVO instance. This poster will present an update of work at WUSM and describe next steps in the project.

Places and Spaces: Mapping Science

Katy Börner, Ph.D. and Elisha F. Hardy (Curators of the Exhibit)

Cyberinfrastructure for Network Science Center School of Library and Information Science, Indiana University, Bloomington, IN 47405

URL: http://scimaps.org/flat/exhibit_info/#1

Abstract

Places & Spaces: Mapping Science is an international science exhibit that aims to inspire cross-disciplinary discussion on how to best track and communicate human activity and scientific progress on a global scale. It has two components: the physical part supports the close inspection of high quality reproductions of maps for display at conferences and education centers; the online counterpart (http://scimaps.org) provides links to a selected series of maps and their makers along with detailed explanations of how these maps work. The exhibit is a 10-year effort. Each year, 10 new maps are added resulting in 100 maps total in 2014. The exhibit was recently on display at the University of Alberta, Edmonton, Alberta, Canada; Stanford University, Stanford, CA; and in the Science Express Train Germany. Currently, Places and Spaces is concluding display at the University of Florida Marsten Science Library. The poster exhibit was recently seen at ECWS 2010, Leiden, The Netherlands, and Eurovis, the 12th annual Visualization Symposium in Bordeaux, June, 2010 and will make its first appearance in Australia at the Joint Conference on Digital Libraries (JCDL) Brisbane, June 2010. The sixth iteration, Science Maps for Scholars, will debut in late Summer 2010.
Laboratree – A Web-based Platform for Team Research Collaboration

Jack H. Pincus, Jamison R. Hemmert, Joy A. Nellis, Brandon, J. Peters, and Sean D. Mooney

Selican Technologies, Inc., Indianapolis, Indiana and The Buck Institute for Age Research, Novato, CA

Slides

Team science projects are collaborative cross-disciplinary research projects organized to solve complex scientific and technical problems. Collaborators within these projects work as virtual teams both inside and between institutions. The Allen Curve\(^1\) shows that interaction between collaborators decreases when they are separated by more than 30 feet. Decreased interaction results in less frequent communication, informal exchange of ideas, document sharing, and data sharing. Cummings and Kiesler\(^2\) concluded collaboration at a distance results in poorer project outcomes.


We are developing Laboratree, as a new web-based research collaboration platform to overcome the challenges of distance collaboration. Laboratree is not a Facebook for scientists or social networking site. It is a secure virtual research environment for project communication, document and data sharing, and real-time informal exchanges of ideas. Laboratree accomplishes document and data sharing with a drag-and-drop folder-based document tree. This feature includes permission-based document sharing, checkout for review and editing, full version control and tracking, and a user configurable metadata system to identify files. Laboratree's project discussion board permits online discussion of experimental results. Its journal (wiki) allows project participants to create interactive and updatable documents. Collaborators can communicate by sending messages to the project team or individual members or hold impromptu chats in real time as the need arises.

Laboratree supports open source and open data standards. We will be implementing application exchange through OpenSocial and incorporating user authentication using OpenID and InCommon. We also plan to add functionality for project and workflow management. VIVO is a natural complement to Laboratree. We plan to use VIVO identities and ontologies to help Laboratree users discover new information about collaborators and make new connections.

Supported by SBIR grant 1R43RR0261155-01 from the National Center for Research Resources.

Using the R Programming Language for VIVO Application Programming

http://semanticommunity.info/Build_VIVO_in_the_Cloud

Updated: Wed, 23 Sep 2015 06:19:59 GMT

Powered by mindtouch
Mike Conlon, PhD.
University of Florida
Gainesville, FL 32653
Track Technical, Application Development, Implementation, Adoption, Collaboration

Slides

Abstract

The R programming language is a popular, open source system for graphics, statistics, and computation. R has outstanding capabilities for processing Extensible Markup Language (XML). VIVO provides its data as linked open data via the Resource Description Framework (RDF), and its XML representation, RDF-Schema. Using R, software developers can rapidly access data hosted by VIVO sites. In this poster we will introduce basic R approaches to consuming VIVO data, including one line programs for accessing VIVO data elements. A series of R programming examples will show how to fetch, process and display data from VIVO. A sample application that builds dynamic email address lists will be presented, as well as a sample social networking application. All code will be provided as open source and posted to both http://vivoweb.org and http://sourceforge.net.

Assessment of Research Impact: A Role for VIVO

Kristi L. Holmes and Cathy C. Sarli
Bernard Becker Medical Library, Washington University School of Medicine
St. Louis, MO 63110

Slides

Abstract

VIVO enables automatic capture of research and scholarship activities of scientists from a variety of reliable sources including local systems, database vendors, and funding agencies. It serves as an aggregated, centralized source of research and scholarship activities that may be used in a variety of ways. One of these ways that VIVO can be utilized is that these activities, "research outputs," serve as a crucial foundation for assessment of research impact beyond traditional means such as citation analysis and use of quantifiable metrics such as number of patents or funded grants, to name a few.

Assessing research impact beyond traditional means in the clinical and biomedical research environment requires tracking research outputs to locate evidence of meaningful health outcomes such as: contribution to the knowledge base; change in understanding of a disease, disorder, or condition; change in practice; change in community health; or change in public law or policy. Starting an "audit trail" using reliable data from a VIVO profile may enhance the assessment process and facilitate greater discovery of indicators of research impact, using a guideline such as the Becker Medical Library Model for Assessment of Research Impact1, 2. The model includes guidance for quantifying and documenting research impact as well as resources for locating evidence of research impact. The six stages of the research process represented by the model include: Research Output; Knowledge Transfer; Clinical Implementation; Collaboration; Community Benefit; and Legislation and Policy Enactment. Having multiple aggregator-contributed research outputs from a single resource such as VIVO offers an ideal starting point to assess research impact across the six stages into a meaningful narrative as a supplement to traditional forms of research assessment.

“It Takes a Village”: Implementation of VIVO at the University of Florida

Amy G. Buhler, MSLS, Engineering Librarian and Associate Chair 1;
Valrie I. Davis, MLS, Agricultural Science Librarian and VIVO National Implementation Coordinator 1;
Sara Russell Gonzalez, PhD, MLIS, Physical Sciences Librarian and UF VIVO Implementation Lead 1;
Margeaux Johnson, MLS, Science and Technology Librarian and Instruction Coordinator 1;
Hannah Norton, MSIS, Reference, Education & Liaison Librarian 2;
Michael Conlon, PhD, Associate Director and Chief Operating Officer, Clinical and Translational Science Institute 3; VIVO Collaboration;

1 University of Florida, Marston Science Library, PO Box 117011, Gainesville, FL 32611-7011
1 University of Florida, Health Science Center Library, PO Box 100206, Gainesville, FL 32610-0206
1 University of Florida, Building 105, PO Box 14283, Gainesville, FL 32604

Slides

Abstract

In 2003 the Cornell University Library staff aligned their expertise with the research and educational missions of their institution and created VIVO – a web-based researcher directory and discovery tool. In 2009, the University of Florida (UF), along with six partner institutions, was awarded $12.2 million by the National Institutes of Health to enhance and expand VIVO into a national network of researchers.

This poster will review the implementation process of VIVO at this large land-grant university at the close of year one of the grant. Conference attendees will be introduced to the institutional characteristics of the University of Florida, the history of the VIVO tool at UF prior to receipt of the grant, the makeup of the UF VIVO Team, and the use of the library-based support model to put this system into operation. In addition, the presentation will detail the current state of the project and a map of future plans for the upcoming year.

Initial comparison of data models in Digital Vita and VIVO

Titus Schleyer, DMD, PhD 1,
William Shirey, MS 2,
Charles Borromeo, BS 2

1 Center for Dental Informatics;
2 Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA

Slides

Digital Vita (DV) is a research networking system developed by the Clinical and Translational Science Institute of the University of Pittsburgh. It integrates CV management functions with academic social networking and basic research
team management and collaboration functions. DV al-lows researchers to manage their complete academic CVs; output
CV information in a variety of formats, such as online profiles and NIH bios-ketches; build their social network through
collaborations on publications and grants automatically, as well as explicit “colleague requests” (equivalent to “friend
requests” in Facebook); and maintain and selectively update multiple versions of NIH biosketches.

VIVO is a national research networking platform funded by the National Center for Research Resources designed to
enable the discovery of research and scholarship across disciplines and institutions. An open source, semantic web
application, it contains researcher interests, activities, and accomplishments. VIVO supports browsing and a search
function which return faceted results for rapid retrieval of desired information. Content in any local VIVO installation may
be maintained manually, brought into VIVO in automated ways from local systems of record, such as human resources,
grants, course, and faculty activity databases, or from database providers such as publication aggregators and funding
agencies.

In order to begin integrating DV with the VIVO platform, we performed a comparative analysis of the data models of the
two systems. Digital Vita’s data model is based on the standard format for the full academic CV at the University of
Pittsburgh. Its information categories include biographical data, education and training, academic appointments, grants
and contracts, awards and honors, publications, presentations, and service activities. It also comprises other
information, such as personal statements (for NIH biosketches) and research interests (for searching and the online
profile). Coauthored items such as publications and presentations are derived from a base class which allows the
system to track collaborators associated with these items. This association allows DV to link collaborators and their
records automatically.

VIVO uses an ontology to represent researchers and their information. Ontologies such as the Fri-end of a Friend
(FOAF) and Bibontology provide the foundation for the VIVO ontology. VIVO adds many classes derived from these
base ontologies to represent people, documents, events and organizations. An example is the sub-classification of a
FOAF person which includes classes for emeritus professor, faculty member, student and others. The VIVO ontology
also adds several top level concepts, such as categories to model biographic information and academic activities, for
instance presentations, teaching and educational background, among many others.

In comparing their respective data models, we noted several areas of congruence and divergence between DV and
VIVO. Both systems model bio-graphical, bibliographic and educational information, and associate people as
collaborators on items such as publications and presentations. VIVO models geographic location, organization and
generic event information, which DV does not include. DV includes support for people to request and track “Colleague”
status with others and while VIVO appears not to. The similarities listed here will be crucial to successful system
integration when data is shared between the two systems. Enough key data elements are present in both systems to
make it likely that system integration at the data level will not pose significant problems.

The results of this project are important for application developers, policy decision makers and others. Application
developers for the VIVO platform must compare the information model of their application to that of VIVO in order to
determine the degree of “data compatibility” between the systems. Policy decision makers must understand which type
of data the respective systems manage in order to determine which applications or combinations thereof can best fulfill
their needs.

i Cole CL, Kanter AS, Cummens M, Vostinar S, Naeymi-Rad F. Using a terminology server and consumer search


**Repurposing VIVO Content with Drupal**

*Miles Worthington*
Interface Designer, Cornell University

Chiang Mai, Thailand

Slides

Abstract

The rich, semantic content within the VIVO system offers Web developers unlimited possibilities in building tools that expand on VIVO's core mission. Through use of Open Linked Data and technologies such as SPARQL, VIVO gives developers a key to content that was formerly locked in locally-defined database schemas – the Deep or Invisible Web.

As one of the largest and most developed open source content management systems (CMS), Drupal provides features and plugins (modules) designed to organize, manipulate, and visualize content in different ways. Whether creating straightforward department or research lab home pages or building complex geographic mapping systems, Drupal is an excellent framework for working with VIVO data.

The major challenge in integrating the two systems lies in mapping VIVO content to fit inside Drupal's relational-database-driven content model. Although there is a comprehensive RDF API available for Drupal 6 as a contributed module, and the forthcoming Drupal 7 adds many core RDF features, the vast majority of Drupal tools are designed to work with content stored in the traditional fashion, inside its SQL database.

My approach for mapping RDF content from VIVO to Drupal involves configuring cross-sections of content as data feeds. These feeds are simply lists of RDF resources (URIs) to be imported and are either entered manually or generated by a SPARQL query. Mappings for each RDF property are configured on a per-feed basis, and at specified intervals feed RDF is fetched via SPARQL query or Linked Data requests.

As the first real-world use of this system, I recently developed a portal site with research summaries from faculty within Cornell's College of Agriculture and Life Sciences (CALS). The goal of the project was to leverage VIVO's web of relationships to create a dynamic faceted browse and search interface that ultimately illustrates the scope of CALS research. Roughly two thousand research summaries from Cornell's VIVO instance were configured for import as feeds, using a Sesame SPARQL endpoint for data access.

The final site serves as a solid proof-of-concept and has helped clarify advantages and disadvantages of a feeds-based approach. Among the Drupal community there a number of projects currently underway that pose exciting possibilities for VIVO-Drupal integration. For the time being, I plan to continue development on the feeds-based approach and intend to release the code as a publicly available Drupal module.

Implementation of VIVO at Ponce School of Medicine –the challenges of a small school setting

Ricardo Espada, Damaris Torres, Michael Vega and Richard Noel

Ponce School of Medicine, Ponce, Puerto Rico

Slides

This poster will provide an overview of the Ponce School of Medicine with an emphasis on the communities most affected by VIVO. It will cover the implementation of VIVO considering issues that we anticipated from the outset as well as unexpected challenges that we have addressed during the first year of implementation. We will also highlight the institutional and personnel resources that have been critical to early adoption, testing and implementation of VIVO.
VIVO Cornell: The Life Cycle of Information

Ellen J. Cramer, PhD. 1, and
Elizabeth (Tobi) Hines, BS 2

1 Special Projects Lead and
2 VIVO Cornell Coordinator Cornell University, Ithaca, NY

Slides

VIVO is a semantic web application that enables the discovery of research and scholarship across disciplines at Cornell. Populated with detailed profiles of faculty and researchers – displaying such information as publications, courses taught, grants received, and professional affiliations – VIVO is a powerful search tool for locating people and information within or across Cornell’s departments, colleges, and campuses. VIVO is both a harvester and distributor of information, with much of the data being collected automatically from Human Resources, Annual Faculty Reporting, the Office of Special Projects, and the Course Database. The data ingested from these authoritative sources is both accurate and current, and reduces the need for manual data input. The result is an integrated and flexible source of publicly visible data available to:

- faculty trying to find collaborators or track competitors
- students trying to locate mentors, advisors, courses, or events
- administrators trying to showcase departmental activities and manage data in one place
- donors or funding agencies trying to keep abreast of research activity at Cornell

VIVO Cornell was originally developed in 2003, since that time it has been successful with ingesting and exposing data. To illustrate VIVO’s role as both a harvester and distributor of information, this poster will graphically display the various data sources being ingested into Cornell University’s VIVO implementation and examples of the re-purposing and re-skinning of this data across the institution, as well as additional ideas for external application development.

VIVO Users, Interface Design, and Evaluation: Lessons Learned

Leslie McIntosh, PhD. 1, Nick Cappadona, MS 2, Manolo Bevia, BS 3, and Ellen J. Cramer, PhD. 4

1 VIVO Evaluation Lead, Washington University in St. Louis, St. Louis, MO
2 User Interface Team Lead, Cornell University, Ithaca, NY
3 Interface Designer, Cornell University, Ithaca, NY and
4 Special Projects Lead, Cornell University, Ithaca, NY

Slides

A significant portion of ongoing and proposed technical innovation related to biomedical research revolves around the goal of facilitating the sharing of data, ancillary information, and resources while enhancing collaboration among researchers across a variety of disciplines. VIVO has been designed to address these needs. To assist in this process and understand how VIVO is achieving these goals, the user interface and evaluation teams are working together to assess and improve VIVO.

Specifically, the teams seek to answer the following questions:
1) How well does the software meet the needs of investigators for finding appropriate people for collaboration and research?
2) How well does the software meet the needs of institutions for learning about their own activities?
3) How accurate and timely are the data at each institution? Is the accuracy related to the techniques used to implement and support the software?
4) What recommendations can be made for improvement?

We will share our mental model diagram, results of the user scenario work, and describe how these along with the user testing results have influenced the development of the application. Examples will be provided of the wireframes and visuals created as part of this process. Additionally, results from the first year-evaluation of VIVO will also be presented.

Development and Application of Subject Knowledge Environment in Chinese Academy of Sciences

Jinxia Huang, Wen Song 1; and

Yi Liu 2

1 National Science Library, Chinese Academy of Sciences, China huangjx@mail.las.ac.cn and songw@mail.las.ac.cn
and

2 Wuhan Branch library of National Science Library, Chinese Academy of Sciences, China liuy@mail.wulib.ac.cn

Keywords: Chinese Academy of Sciences (CAS); Subject Knowledge Environment (SKE); Knowledge Organization System (KOS); Ontology; Vitro; Application

Slides

Abstract

In order to provide the services on knowledge discovery and help the scientists in Chinese Academy of Sciences (CAS) communicate and share knowledge, National Science Library (NSL) of CAS has developed the Subject Knowledge Environment (SKE) that uses an entity-relationship ontology model as the Knowledge Organization System (KOS) to integrate and organize different kinds of scientific information, such as people, organization, project, academic event, scientific equipment, funding opportunity, literature resources.

With the changes in current research life-cycle and the scientific workflow, the needs from the scientists are also being changed in which the services on the communication and dissemination on scientific information and knowledge are put forward. SKE aims to work as a domain-focused academic community for the scientists to realize the relationships between resources and knowledge reasoning, to support the librarians to provide the real-time and high-valued knowledge services, to support the researchers to organize and manage the knowledge resources, and to support the knowledge mining and discovery.

Some related work focused on the literature reviews on Virtual Research Environment, Semantic Web, Knowledge Organization System, and the surveys on Social Network Services was completed firstly, and then SKE system architecture and SKE ontologies were designed. Based on the open resource software Vitro which was developed by Cornell University Library and the customization on the system, SKE system platform was constructed quickly and now is being applied in some academic groups in CAS, such as the institute’s knowledge environment, the project’s knowledge environment and the research group’s knowledge environment. The ontology management, data mining and the widely application on SKE will be the future work.
The poster reports the functionality, the system architecture, ontology generation, main services and functions, and applications on SKE.

**Comparative Matrix of Research Networking Tools**

**Holly Falk-Krzesinski, PhD** 1; **Pamela L. Shaw, MS, MSLIS** 2; **Laura Wimbiscus-Yoon, MS** 3

1 Research Team Support (RTS) Northwestern University Clinical and Translational Sciences (NUCATS) Institute, 750 N Lake Shore Drive, 11th Floor, Chicago, IL 60611 USA

2 Galter Health Sciences Library, Northwestern University Feinberg School of Medicine, 303E Chicago Avenue, Chicago, IL 60611 USA

3 Northwestern University Biomedical Informatics Center (NUBIC), NUCATS Institute, 750N Lake Shore Drive, 11th Floor, Chicago, IL 60611 USA

Slides

Northwestern University is geographically divided into two campuses: a main, undergraduate campus in Evanston, IL that houses seven schools, and a downtown Chicago campus that is home to three schools. Moreover, the University has three geographically dispersed medical/hospital affiliates. Translational research is performed at both campuses and in conjunction with medical partners, but researchers are generally not aware of expertise, resources, and facilities available at the location that is not their own—even of the research landscape beyond their own school or department! The Northwestern University Clinical and Translational Sciences (NUCATS) Institute recognized the need for a knowledge management system of the University’s entire research enterprise, to gain access to the full spectrum of research resources and with the important goal of encouraging more research collaboration among faculty across all campuses/schools and affiliated medical facilities. The NUCATS Director of Research Team Support launched a project to collect information for a comparative matrix of available research networking tools. This effort involved members of the Galter Health Sciences Library and the Northwestern University Biomedical Informatics Center (NUBIC) at the NUCATS Institute (an additional 17 units at the University and medical affiliates were engaged at the product demonstration stage of the process). This team identified 25 open source and licensed applications for research networking. Sixteen factors were identified for comparison of the tools, including time required for implementation, interoperability with internal and external data sources, and data infeed functionality. Though the project was undertaken to identify the best research networking tool solution for Northwestern University, the team was encouraged to share their findings with other institutions undertaking a similar investigation of research knowledge management systems. The completed comparison provides a glimpse of one institution’s findings in the search of an effective research networking application.

**Virtual Appliance: releasing without an executable**

**Christopher Barnes, Christopher Haines, Yang Li, Narayan Raum, Dale Scheppler, Nicholas Skaggs, and Stephen V. Williams**

University of Florida Gainesville, FL USA

Slides

Abstract

Virtual machines allow network administrators to operate multiple systems on a single physical server. These systems act as an thin layer between the OS and the hardware, allowing multiple OS instances to be run independently. The consolidation of resources allows for the creation of redundant systems at a much lower cost. On an individual users
machine virtual machines allow the users to run applications tailored for a specific system on their home environment through a common client. Snapshots provide a safety net, so users and administrators alike can feel more relaxed about making modifications to the installation. If something does not work, then the snapshot can be used to return to the original configuration. Administrators of mission critical applications can launch second copies of a system, perform upgrades and test the stability all the while, leaving the original running.

Virtual Machine Appliances allow software packagers to release ready-to-use versions of their software in a sandboxed environment. These appliances are pre-configured installations of all required software and operating environment. A user simply has to power on the virtual machine and they have a fully functional system. With the simplicity of getting the system running and restoring it to a previous snapshot, appliances have a wide variety of uses. From providing a safe learning environment, to marketing a software product, demoing the system with sample data, production systems, and test platforms; virtual machine appliances provide a flexible means to release your software in ready to use bundles, easy to install and even easier to throw away.

Creating a collaborative research network for scientists –the example of Mendeley

Jan Reichelt and Victor Henning
Mendeley
London, United Kingdom

Keywords: Collaboration, social networking, crowdsourcing, linked open data, research trends

Slides

Abstract

This abstract proposes a 10-minute presentation of Mendeley at the VIVO conference 2010 with the aim of sharing knowledge and “lessons learned” in order to improve collaboration between scientists. Mendeley is a research workflow and collaboration tool which crowdsources real-time research trend information and semantic annotations of research papers in a central data store, thereby creating a “social research network” that emerges based on research data. We describe how Mendeley’s model can overcome barriers for collaboration by turning research papers into social objects and making academic data publicly available via an open API.

Central to the success of Mendeley has been the creation of a tool that works for the researcher without the requirement of being part of an explicit social network. Mendeley automatically extracts metadata from research papers, and allows a researcher to annotate, tag and organize their research collection. The tool integrates with the paper writing workflow and provides advanced collaboration options, thus significantly improving researchers’ productivity. By anonymously aggregating usage data, Mendeley enables the emergence of social metrics and real-time usage stats on top of the articles’ abstract metadata. In this way a social network of collaborators, and people genuinely interested in content, emerges. By building this research network around the article as the social object, a social layer of direct relevance to academia emerges.

Within 18 months, Mendeley’s userbase has grown to more than 400,000 users, and the database has grown to more than 30 Million entries, setting Mendeley on track to become the largest open academic database by the end of 2010. Information in the database is accessible directly via Mendeley, or alternatively via Mendeley’s open API, which attracts researchers to create value-added applications on top of Mendeley, thus further enriching the data and incentivizing collaboration. Significant challenges have had to be overcome in creating a tool that is stable for hundreds of thousands of users, both technically and conceptually. Currently, additional efforts go into data mining activities, such as article and author name disambiguation, entity extraction, recommendation engines, and enriching the existing network with semantic information.
Drawing Organizational Charts Using VIVO

Alex Rockwell

University of Florida Gainesville, FL

VIVO provides complete information on organizational structures of institutions. Each organization object in VIVO has parent and child organizations. Starting at any particular organization, one can easily use a simple recursion algorithm to traverse the organizations that report up to the starting point. If the starting point is the institution “root”, one can produce an organizational chart for the entire organization. Using Ruby and some open source extensions, simple software has been developed to draw pictures of organizations. Code, algorithm, commentary and sample output will be presented. All code will be available as open source at http://sourceforge.net

VIVO Development Road Map Poster

Jon Corson-Rikert, BA 1; Christopher P. Barnes 2; Micah Linnemeier, BS 3; Stephen Williams 2; Narayan Raum 2; Nick Cappadona, MS 1; Brian Caruso, BA 1; Brian Lowe, BA 1

1 Cornell University, Ithaca, NY and
2 University of Florida, Gainesville, FL
3 Indiana University, Bloomington, IN

The VIVO Project has created a development road map charting improvements and new features for the VIVO application spanning the life of the grant, through August of 2011. This road map was written for internal project use by the development teams and as such is not an ideal document for presentation to the full project team, interested developers outside the project, or potential adopters of VIVO.

As one step toward broader consideration of the development plans for the VIVO project, significant efforts have been made to align the development road map with the recently-completed user scenario process to assess whether, how, and to what extent the most important and highest priority scenarios can be implemented. This review has produced additional documentation, but again not in an optimal format to support high-level discussion of the development directions of project.

The Development Road Map Poster will be compiled to communicate in broad strokes the development accomplishments to date and plans for ongoing development within the scope of the VIVO project. Content will be drawn from the existing road map documentation and timelines contributed by the leads of each of the development teams. The focus will be on function and features rather than technical implementation detail. We anticipate the poster will facilitate broad understanding of development goals, data ingest and export, and opportunities for collaborative development extending VIVO for specific purposes and/or audiences.

The poster will be formatted at a level of detail and with sufficient graphic clarity to promote review and one-on-one or small group discussion during the poster session. Development team leads will be encouraged to be available for informal discussion, and the poster and discussion will complement the previously accepted full-length paper presentation by many of the same authors, entitled “VIVO Development Overview.”
Weill Cornell Medical College VIVO Implementation

Kenneth Lee, Dan Dickinson and Curtis Cole

Weill Cornell Medical College, New York, NY, United States

Slides

The mission of the Medical College is to provide the finest education possible for medical students and students pursuing advanced degrees in the biomedical sciences, to conduct research at the cutting edge of knowledge, to improve the health care of the nation and the world, and to provide the highest quality of clinical care to the community.

For many researchers the geographical and organizational confines of a department, college, or even a single university bear very little relevance to the scope of their research or the pool of colleagues they may seek for collaboration. Researchers are often left to find their own paths to discover current activities and active researchers in their field and beyond, usually by a combination of personal connection, disciplinary knowledge, and fortuitous discovery through search engines, leaving those who have yet to develop their own network of personal contacts at a significant disadvantage.

Weill Cornell’s participation in this NIH U24 grant provides the opportunity to share our internal expertise in faculty profile management as well as expand our internal effort through a national collaboration using cutting edge technology beyond an isolated effort. We have made great progress during the first year of the grant, some of our accomplishments include identifying authoritative data sources, defining preliminary ontology and establishing a pilot department. Some of our goals for the next year include integrating VIVO into the WCMC culture and then rolling out VIVO to our partner institutions.