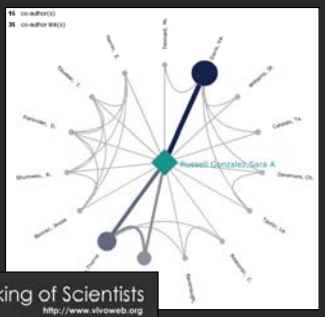
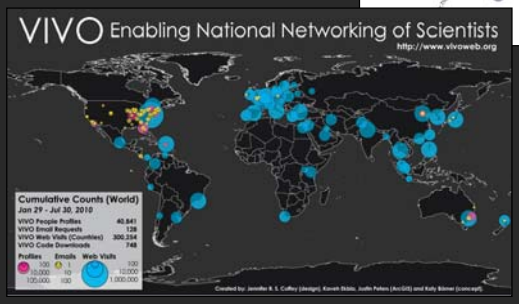
VIVO and VIVO@IU

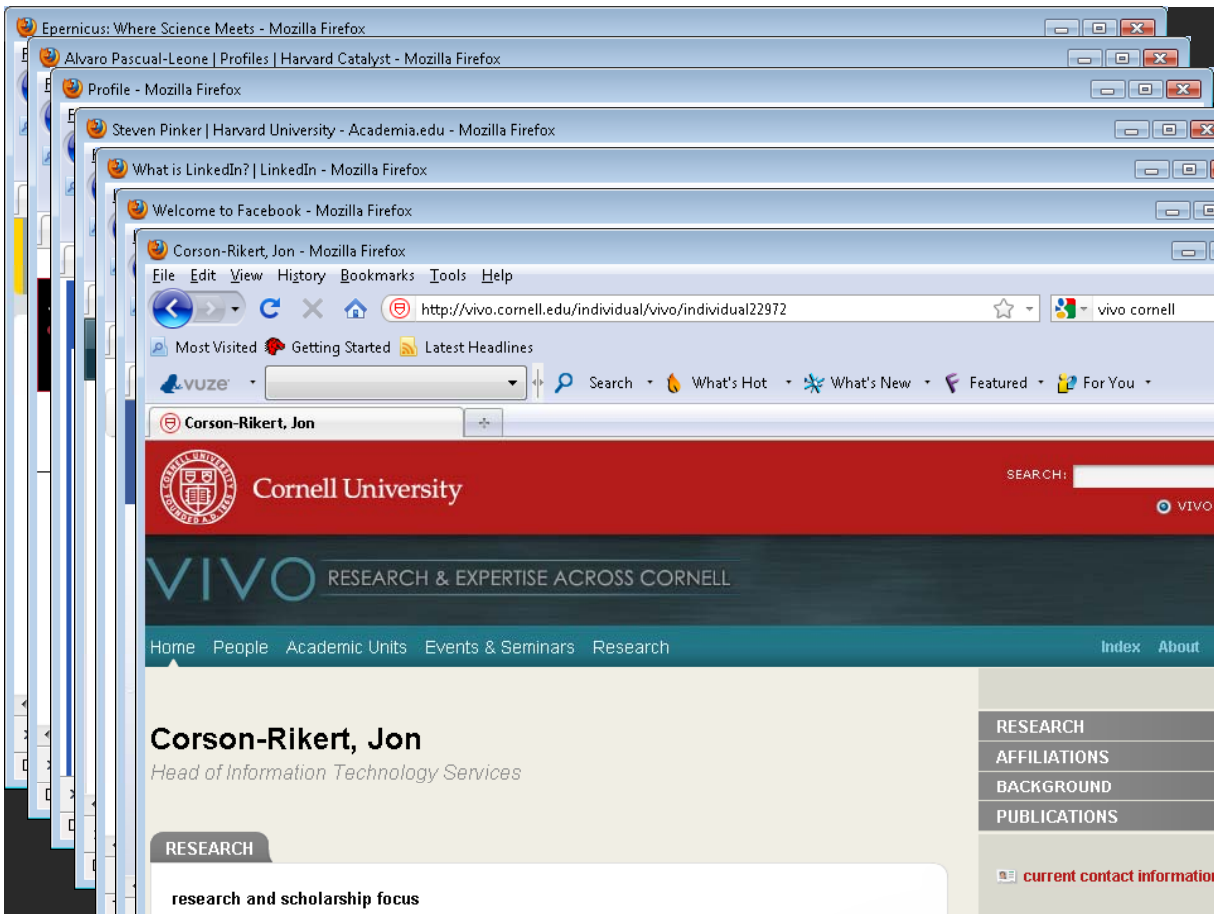
Dr. Katy Börner
Cyberinfrastructure for Network Science Center, Director
Information Visualization Laboratory, Director
School of Library and Information Science
Indiana University, Bloomington, IN
katy@indiana.edu

With special thanks to the members at the Cyberinfrastructure for Network Science Center and the VIVO team.

UITS Presentation
September 17, 2010





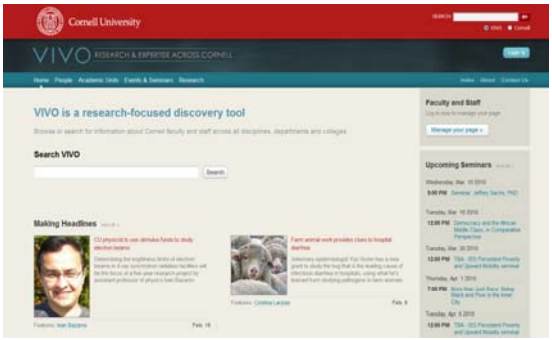


VIVO: A Semantic Approach to Creating a National Network of Researchers (http://vivoweb.org)

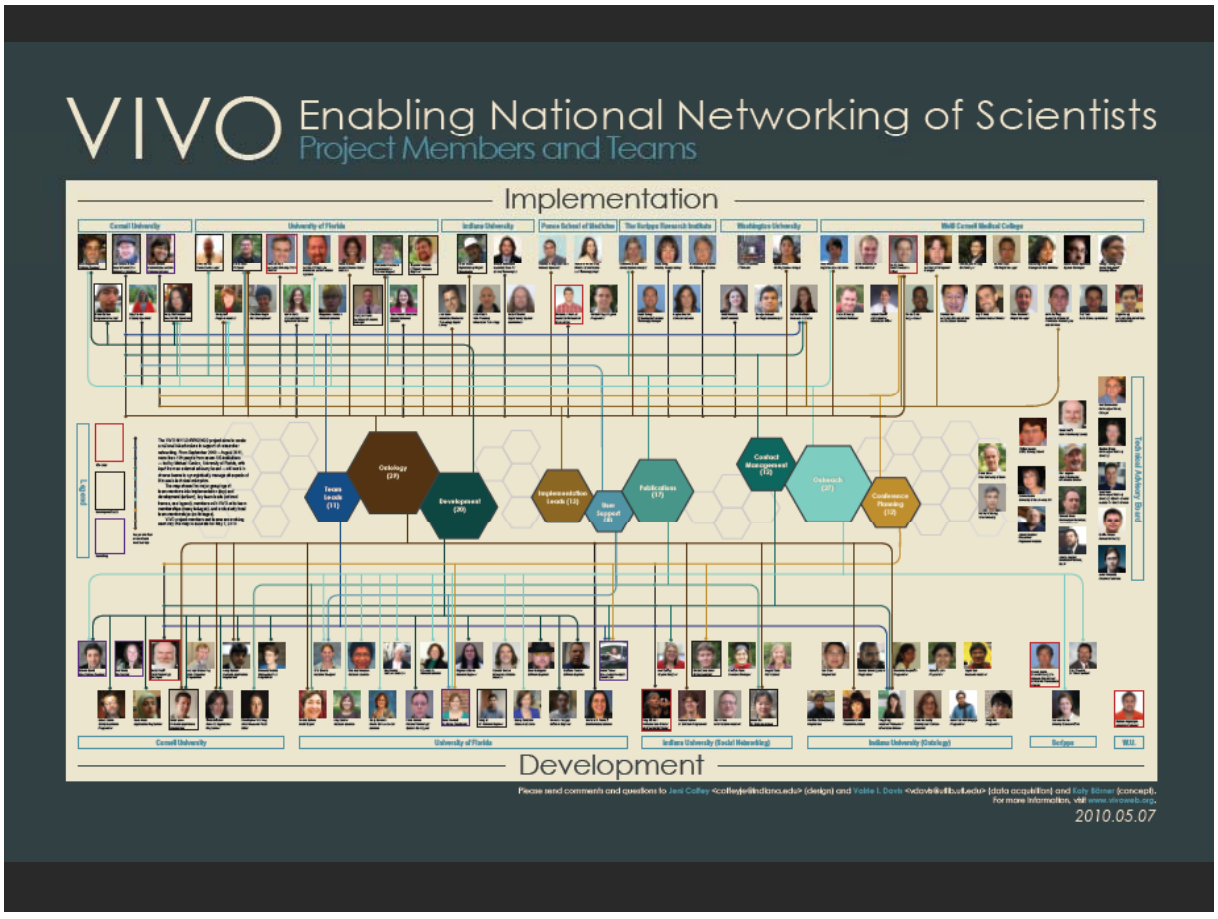
- Semantic web application and ontology editor originally developed at Cornell U.
- Integrates research and scholarship info from systems of record across institution(s).
- Facilitates research discovery and cross-disciplinary collaboration.

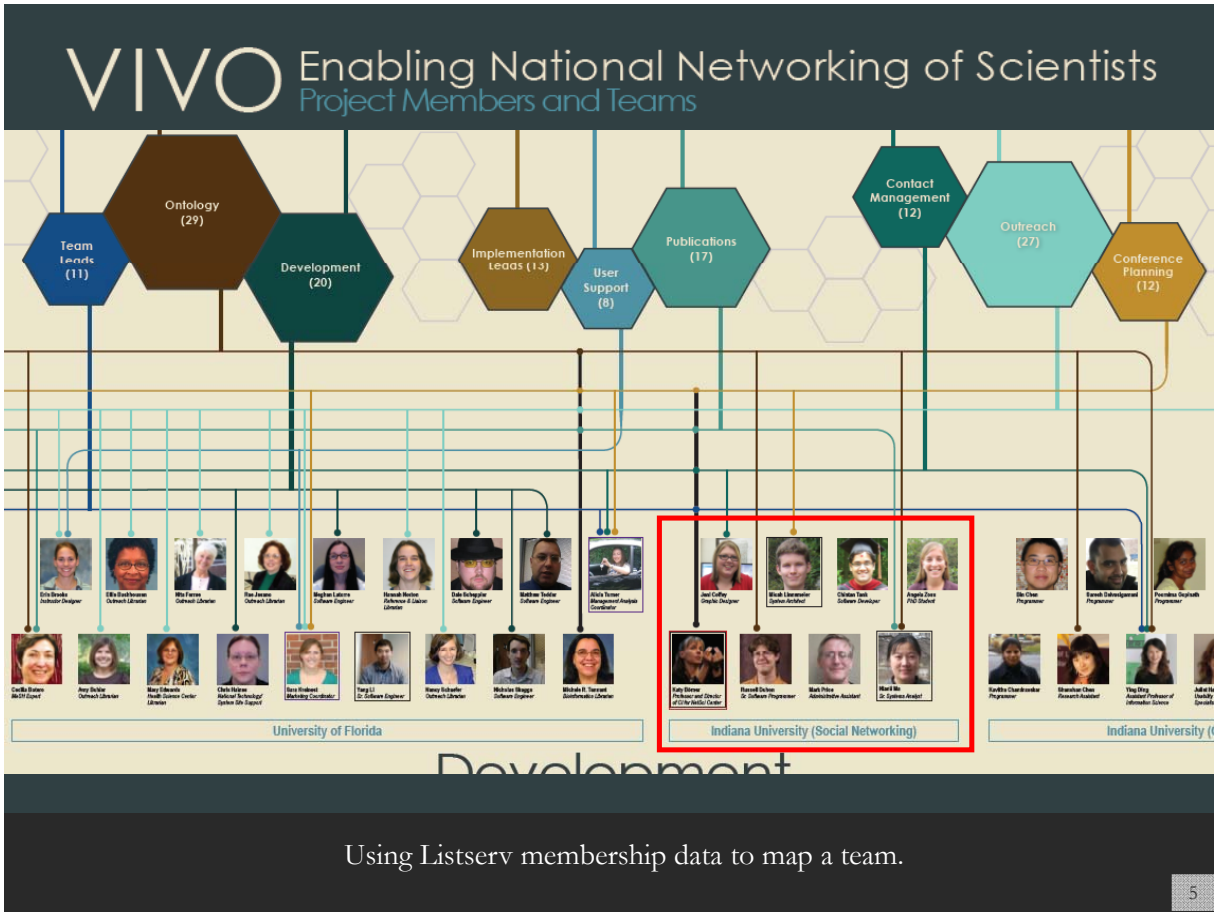
Soon:

- Simplify reporting tasks, e.g., generate biosketch, department report.



Cornell University: Dean Krafft (Cornell PI), Manolo Bevia, Jim Blake, Nick Cappadona, Brian Caruso, Jon Corson-Rikert, Elly Cramer, Medha Devare, John Fereira, Brian Lowe, Stella Mitchell, Holly Mistlebauer, Anup Sawant, Christopher Westling, Rebecca Younes. **University of Florida:** Mike Conlon (VIVO and UF PI), Cecilia Botero, Kerry Britt, Erin Brooks, Amy Buhler, Ellie Bushhousen, Chris Case, Valrie Davis, Nita Ferree, Chris Haines, Rae Jesano, Margeaux Johnson, Sara Kreinest, Yang Li, Paula Markes, Sara Russell Gonzalez, Alexander Rockwell, Nancy Schaefer, Michele R. Tennant, George Hack, Chris Barnes, Narayan Raum, Brenda Stevens, Alicia Turner, Stephen Williams. **Indiana University**: Katy Borner (IU PI), William Barnett, Shanshan Chen, Ying Ding, Russell Duhon, Jon Dunn, Micah Linnemeier, Nianli Ma, Robert McDonald, Barbara Ann O'Leary, Mark Price, Yuyin Sun, Alan Walsh, Brian Wheeler, Angela Zoss. **Ponce School of Medicine:** Richard Noel (Ponce PI), Ricardo Espada, Damaris Torres. **The Scripps Research Institute:** Gerald Joyce (Scripps PI), Greg Dunlap, Catherine Dunn, Brant Kelley, Paula King, Angela Murrell, Barbara Noble, Cary Thomas, Michaeleen Trimarchi. **Washington University, St. Louis**: Rakesh Nagarajan (WUSTL PI), Kristi L. Holmes, Sunita B. Koul, Leslie D. McIntosh. **Weill Cornell Medical College:** Curtis Cole (Weill PI), Paul Albert, Victor Brodsky, Adam Cheriff, Oscar Cruz, Dan Dickinson, Chris Huang, Itay Klaz, Peter Michelini, Grace Migliorisi, John Ruffing, Jason Specland, Tru Tran, Jesse Turner, Vinay Varughese.





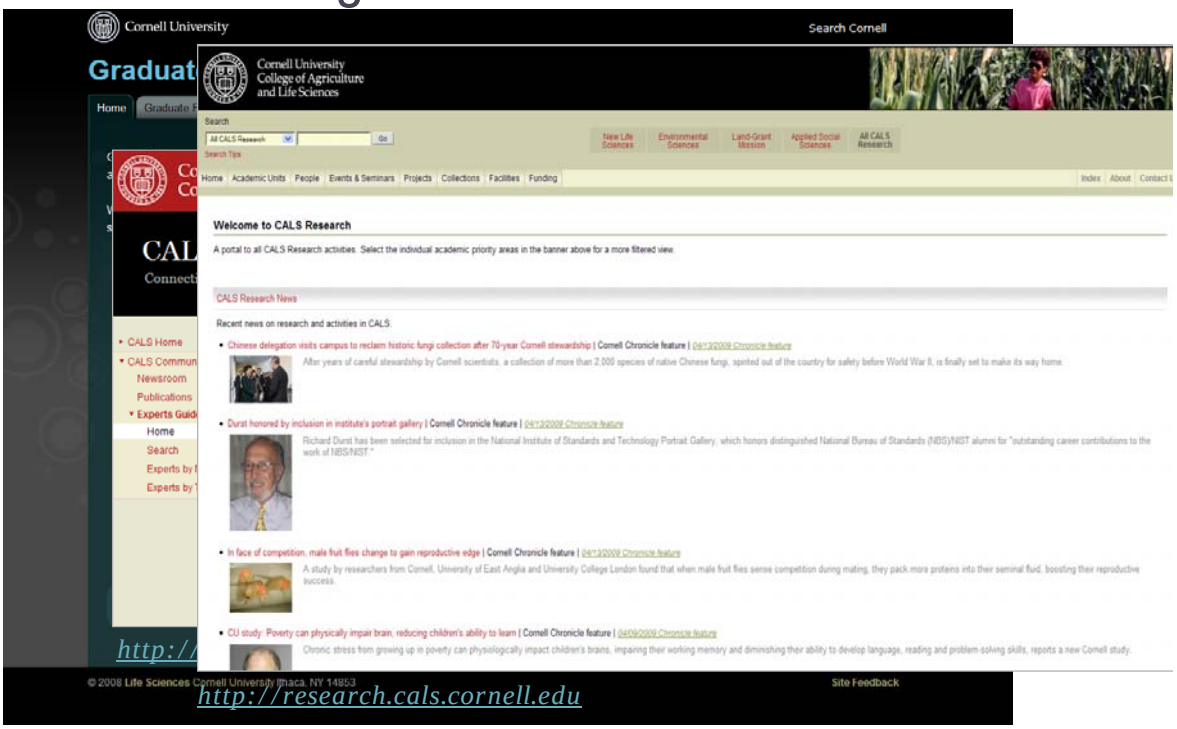
Using Listserv membership data to map a team.

VIVO Enabling a National Network of Scientists

VIVO Users and Needs

- Faculty/Researchers
  - Customize profile created via feeds; find potential collaborators, "people like me"; discovery via high search rankings; info on activity of colleagues…
- Students
  - Create profiles; easily find mentors + collaborators; locate facilities, events, funding opportunities…
- Administrators
  - Quickly find cross-disciplinary expertise (research area; geography); centralize public data from diverse sources; easily repurpose information for consumers; improve faculty collaboration within or across departments and institutions…
- Funding, donor, legislative agencies
  - Discover projects, grants, expertise (e.g. for review panels; targets for funding)…
- General public
  - Find expertise, learn about research in a region/institution…

VIVO Web Pages



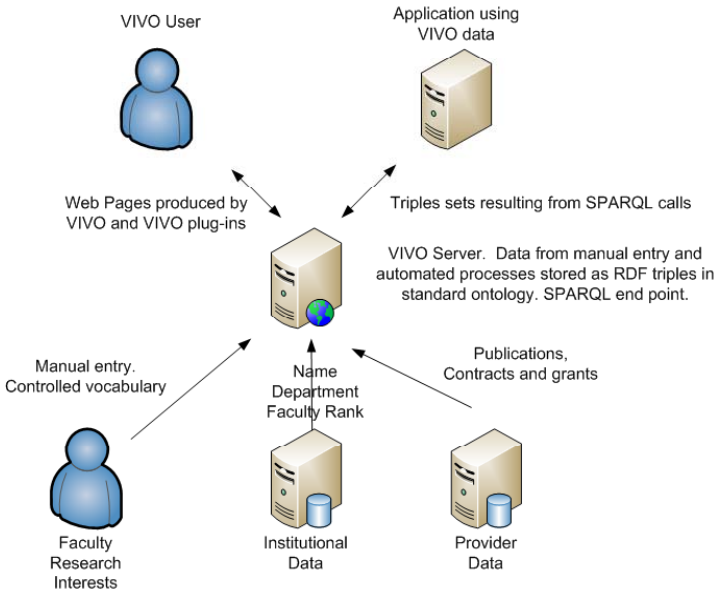
http://gradeducation.lifesciences.cornell.edu

VIVO Data Providers & Users

- Eagle-i ("enabling resource discovery" U24 award)
- Federal agencies – NIH (NIH RePORTER), NSF, USDA, …
- Search Providers – Google, Bing, Yahoo, …
- Professional Societies – AAAS, …
- Publishers/vendors – PubMed, Elsevier, Collexis, ISI…
- Semantic Web community – DERI, …
- Consortia of schools – SURA, CTSA…
- Producers, consumers of semantic web-compliant data

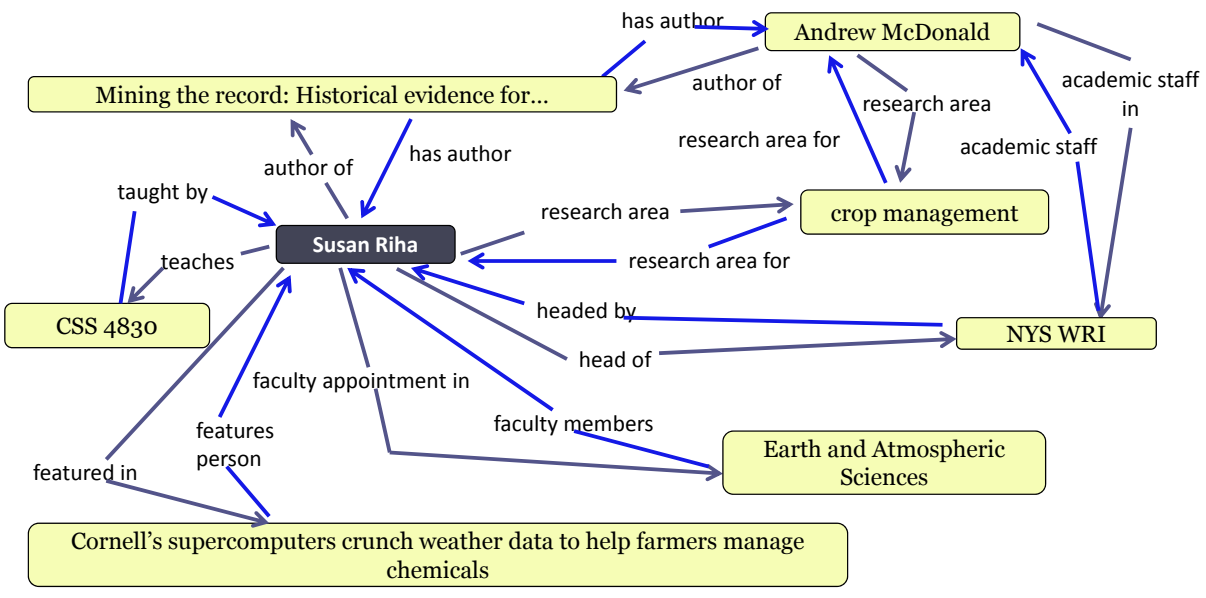
Institutional Architecture

- Three sources of VIVO information
  - User data
  - Institutional data
  - Provider data
- Two formats for output
  - Web Pages for users
  - Resource Description Framework for applications



Data Representation using RDF Triples

Detailed relationships for a researcher at Cornell U.
Open source code (BSD) and ontology available at http://vivoweb.org.



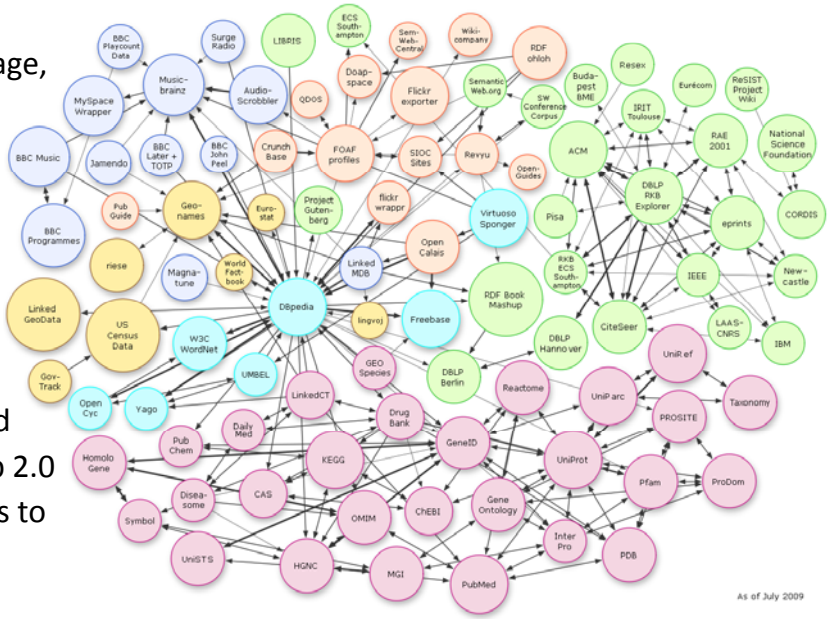
VIVO & Linked Open Data

2010 National VIVO Conference August 12&13, NYC
http://conferences.dce.ufl.edu/vivo

VIVO makes high coverage, high quality data from systems of record

- available online
- for free, and
- in machine readable format.

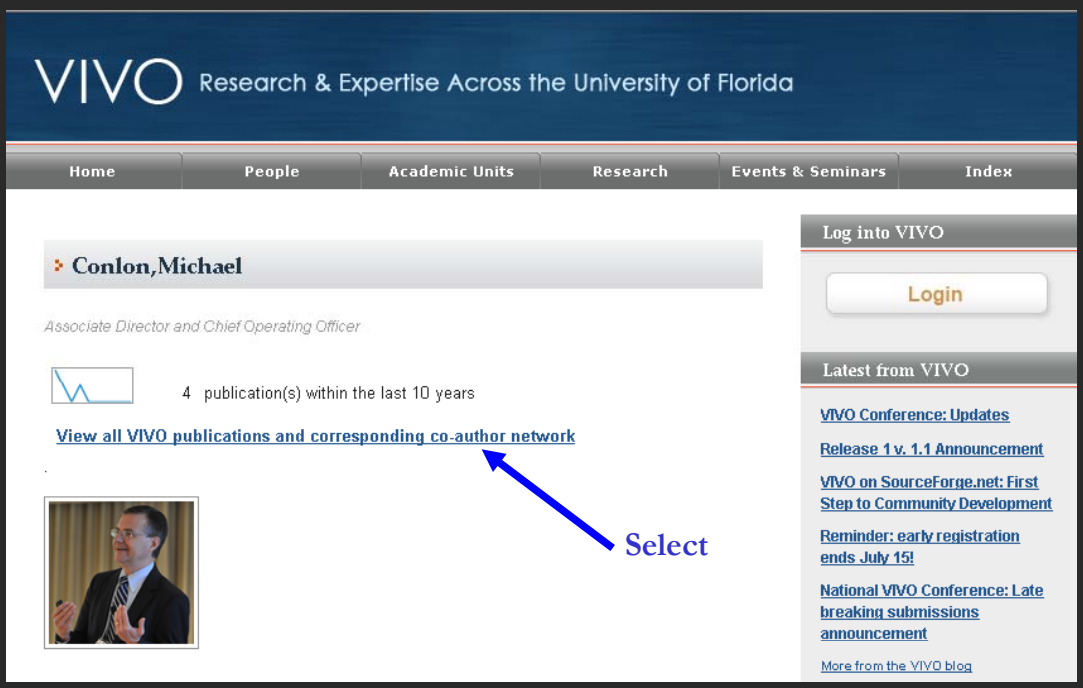
VIVO ontology is aligned with many existing Web 2.0 and scholarly ontologies to ease interoperability.



http://www4.wiwiss.fu-berlin.de/bizer/pub/lod-datasets_2009-07-14_colored.png

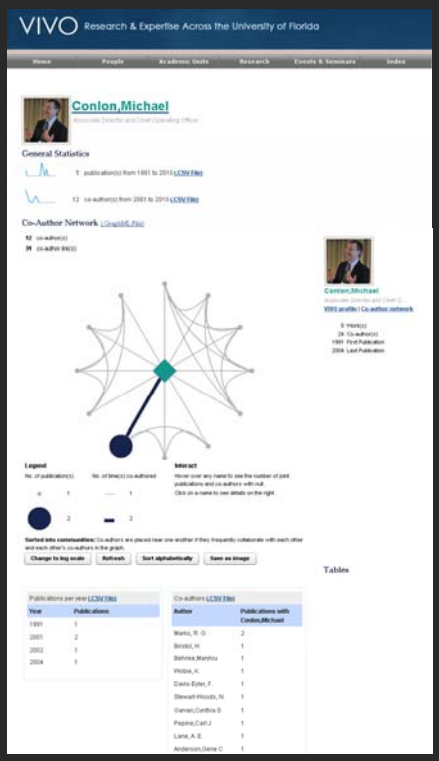
VIVO ENABLING NATIONAL NETWORKING OF SCIENTISTS

VIVO Release 1 v. 1.1: Individual Level Co-Author Visualization



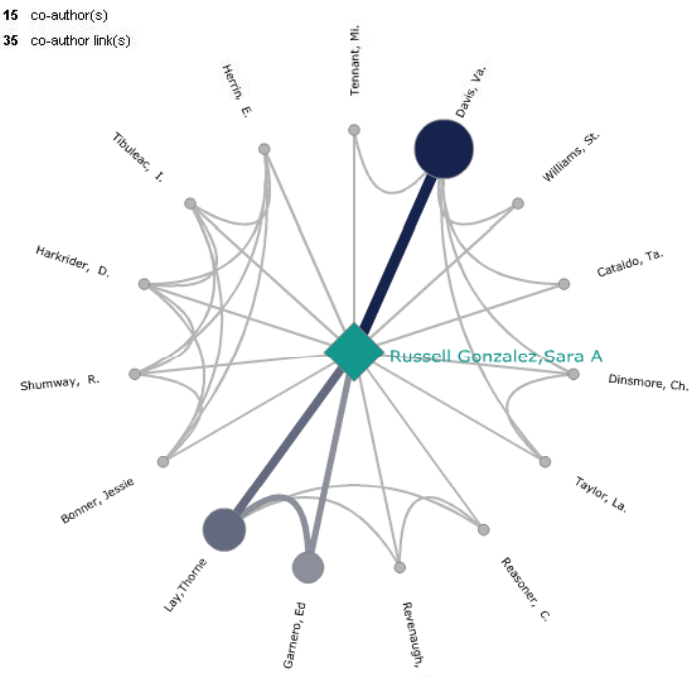
http://vivo.ufl.edu/display/n2556

VIVO Release 1 v. 1.1: Individual Level Co-Author Visualization



Co-Author Network (GraphML File)

15 co-author(s)
35 co-author link(s)

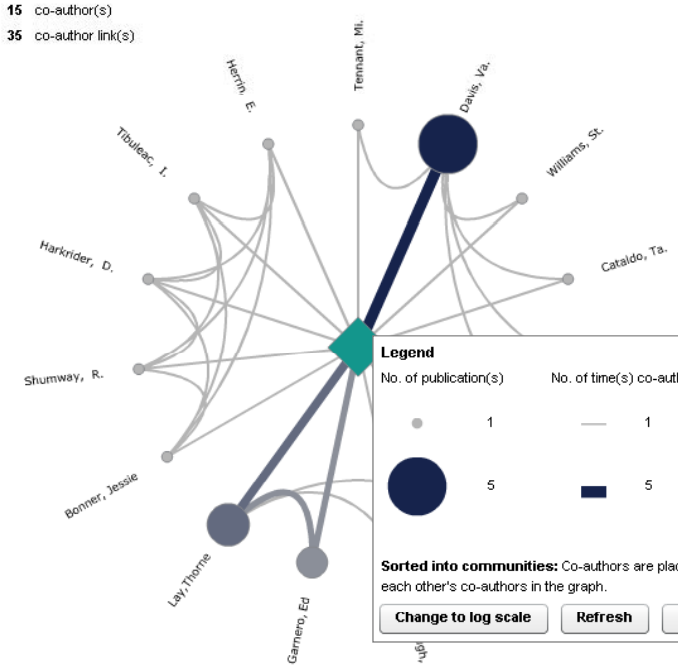


Russell Gonzalez,Sara A
Physical Sciences Librarian
VIVO profile | Co-author network

9 Publication(s)
15 Co-author(s)
1998 First Publication
2010 Last Publication

Co-Author Network (GraphML File)

15 co-author(s)
35 co-author link(s)



Russell Gonzalez,Sara A
Physical Sciences Librarian
VIVO profile | Co-author network

9 Publication(s)
15 Co-author(s)
1998 First Publication
2010 Last Publication

Legend

No. of publication(s): 1, 5
No. of time(s) co-authored: 1, 5

Interact

Hover over any name to see the number of joint publications and co-authors with Russell Gonzalez,Sara A.
Click on a name to see details on the right.

Thresholding

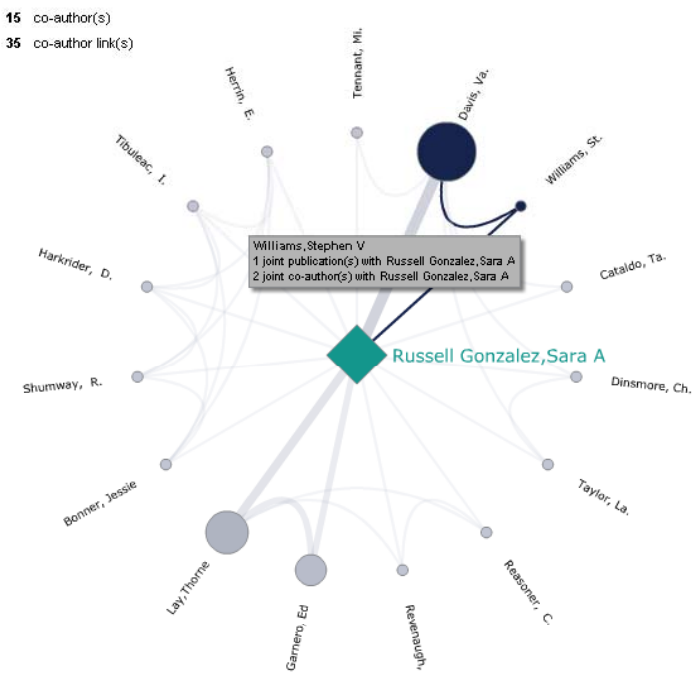
Only people that co-authored more than 0 paper(s) with Russell Gonzalez,Sara A are shown.
0 out of 0 co-author(s) are shown.

Sorted into communities: Co-authors are placed near one another if they frequently collaborate with each other and each other's co-authors in the graph.

Change to log scale | Refresh | Sort alphabetically | Save as image

Co-Author Network (GraphML File)

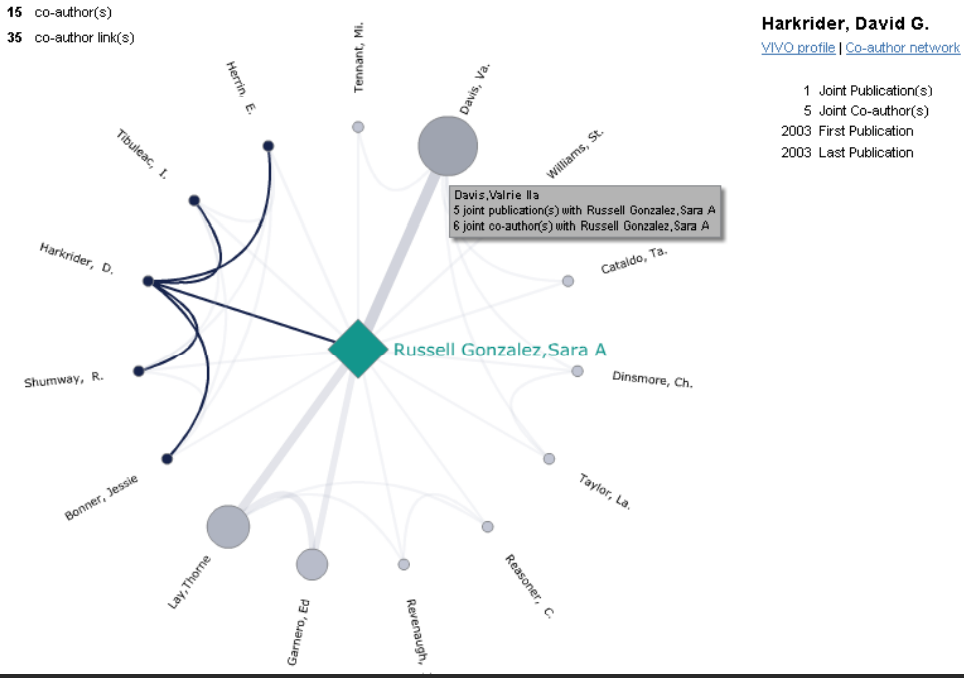
15 co-author(s)
35 co-author link(s)

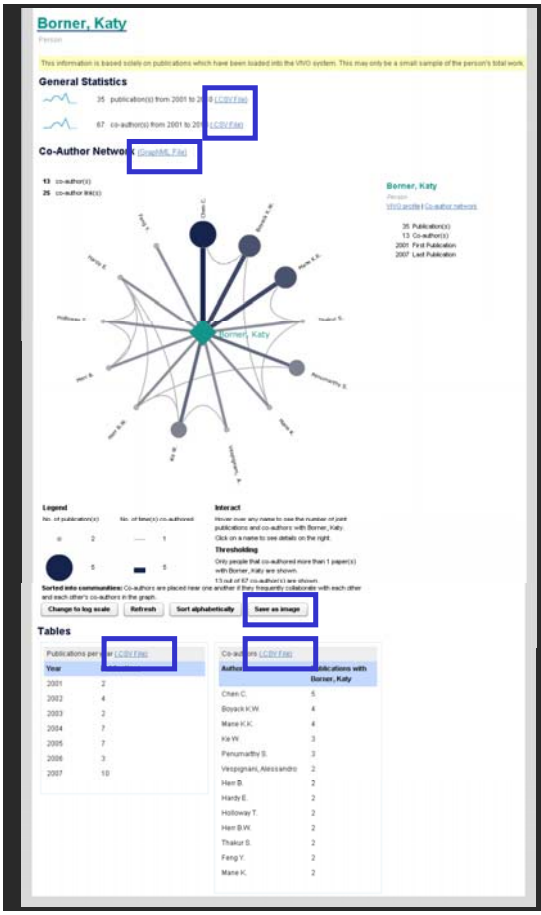


Williams,Stephen V
Software Engineer
VIVO profile | Co-author network

1 Joint Publication(s)
2 Joint Co-author(s)
2010 First Publication
2010 Last Publication

Co-Author Network (GraphML File)





Download Data

General Statistics

- 36 publication(s) from 2001 to 2010 (.CSV File)
- 80 co-author(s) from 2001 to 2010 (.CSV File)

Co-Author Network

(GraphML File)

Save as Image (.PNG file)

Tables

- Publications per year (.CSV File)
- Co-authors (.CSV File)

http://vivo-vis.slis.indiana.edu/vivo1/visualization?uri=http%3A%2F%2Fvivoweb.org%2Fontology%2Fcore%2FPerson72c&vis=person_level&render_mode=standalone

36 publication(s) from 2001 to 2010 (.CSV File)

Year	Publications
2001	2
2002	4
2003	2
2004	7
2005	7
2006	3
2007	10
2010	1

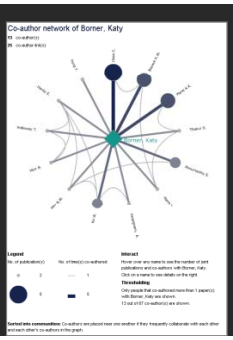
80 co-author(s) from 2001 to 2010 (.CSV File)

Year	Count	Co-Author(s)
2001	1	Chen C.
2002	3	Chen C.; McMahon T.; Feng Y.
2003	2	Chen C.; Boyack K.W.
2004	17	Sengupta A.; Penumarthy S.; Thakur S.; Sooriamurthi R.; Maru J.T.; Shiffrin R.M.; Mane K.; Moor K.A.;

Co-author network (GraphML File)

```
<?xml version="1.0" encoding="UTF-8"?>
    <graphml xmlns="http://graphml.graphdrawing.org/xmlns"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://graphml.graphdrawing.org/xmlns
  http://graphml.graphdrawing.org/xmlns/1.0/graphml.xsd">
<key id="label" for="node" attr.name="label" attr.type="string" />
<key id="number_of_authored_works" for="node" attr.name="number_of_authored_works" attr.type="int" />
<key id="num_unknown_publication" for="node" attr.name="num_unknown_publication" attr.type="int" />
<key id="num_latest_publication" for="node" attr.name="num_latest_publication" attr.type="int" />
<key id="latest_publication" for="node" attr.name="latest_publication" attr.type="int" />
<key id="profile_url" for="node" attr.name="profile_url" attr.type="string" />
```

Save as Image (.PNG file)

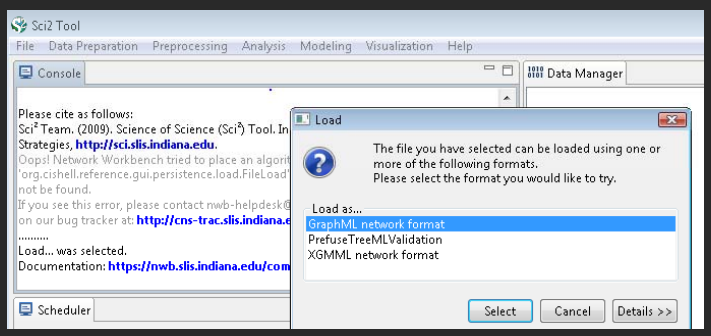


Publications per year (.CSV File), see top file.

Co-authors (.CSV File)

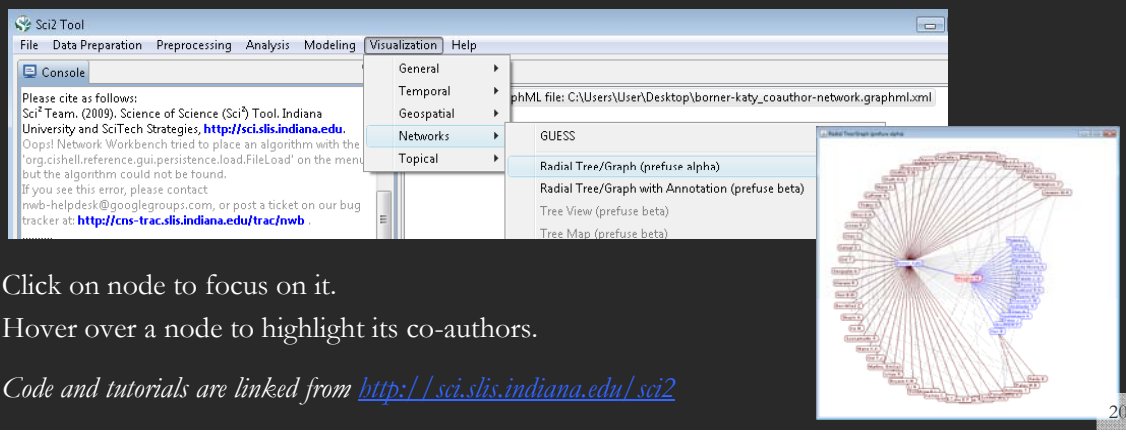
Co-Author	Count
Andrienko G.	1
Andrienko N.	1
Ben-Miled Z.	1
Blackwell A.	1
Boyack K.W.	4
Bozicevic M.	1
Brodbeck D.	1
Burkhard R.A.	1
Chen C.	5

Run Sci2 Tool and Load Co-Author Network (GraphML File)



Network Analysis Toolkit
Nodes: 81
Edges: 390

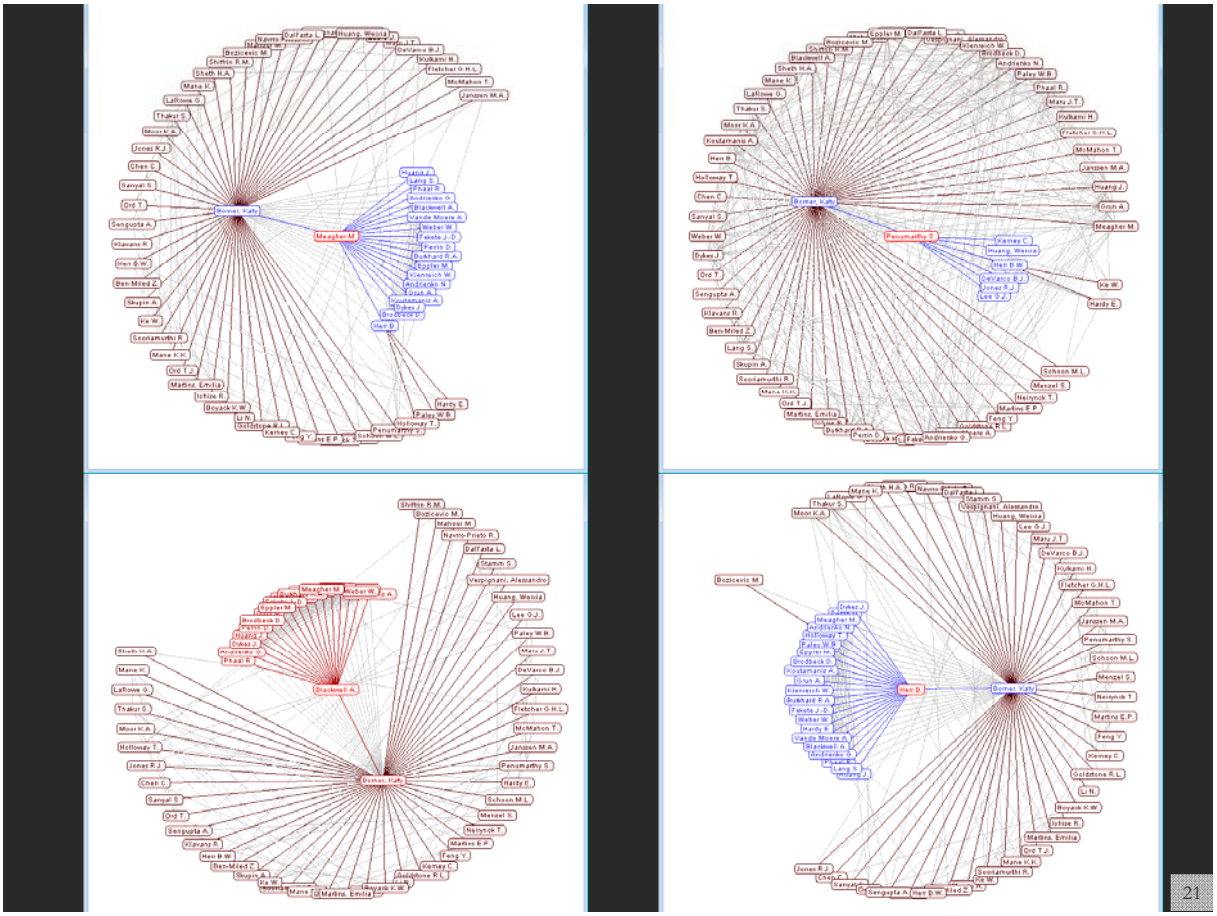
Visualize the file using Radial Graph layout.



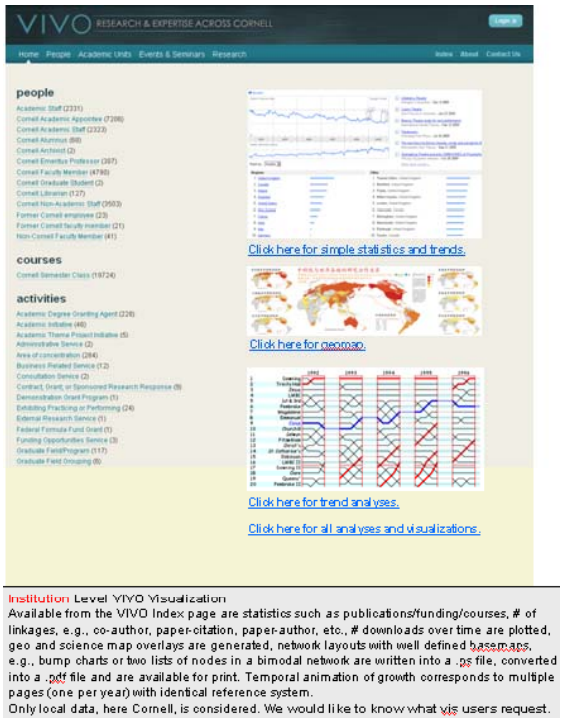
Click on node to focus on it.

Hover over a node to highlight its co-authors.

Code and tutorials are linked from http://sci.slis.indiana.edu/sci2



VIVO Institution Level Visualizations



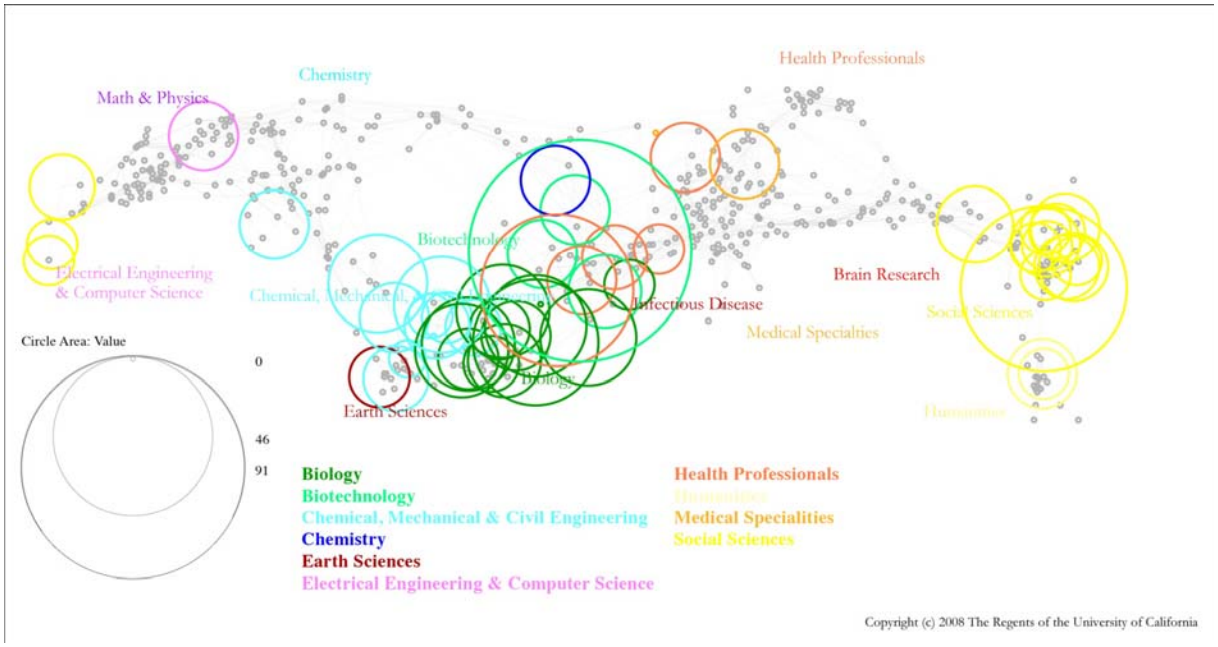
Institution level visualization will be available from the VIVO Index page and comprise statistics such as

- publications/funding/courses,
- # of linkages, e.g., co-author,
- paper-citation, paper-author, etc.,
- # downloads over time are plotted.

Geospatial and science map overlays as well as network layouts with well defined base maps, e.g., two lists of nodes in a bimodal network will be written into a PDF file for viewing and printing.

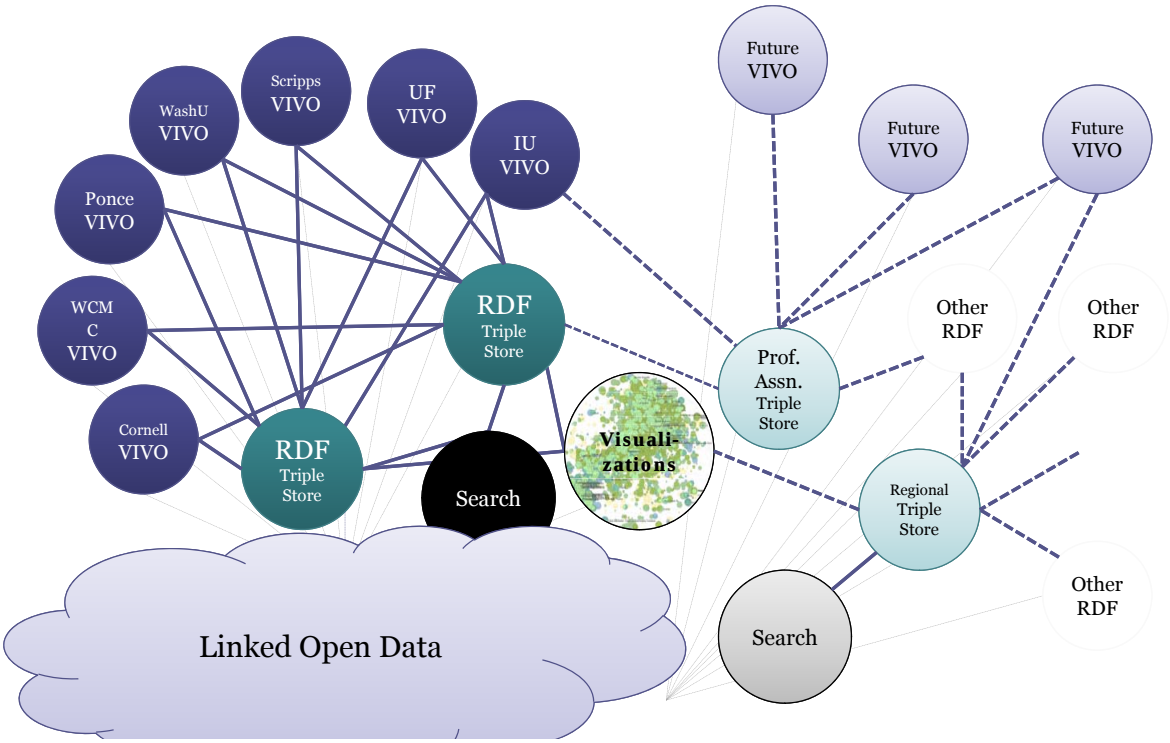
Temporal animation of growth corresponds to multiple pages (one per year) with identical reference system.

(Generated using dummy data. The values shown here are not real).

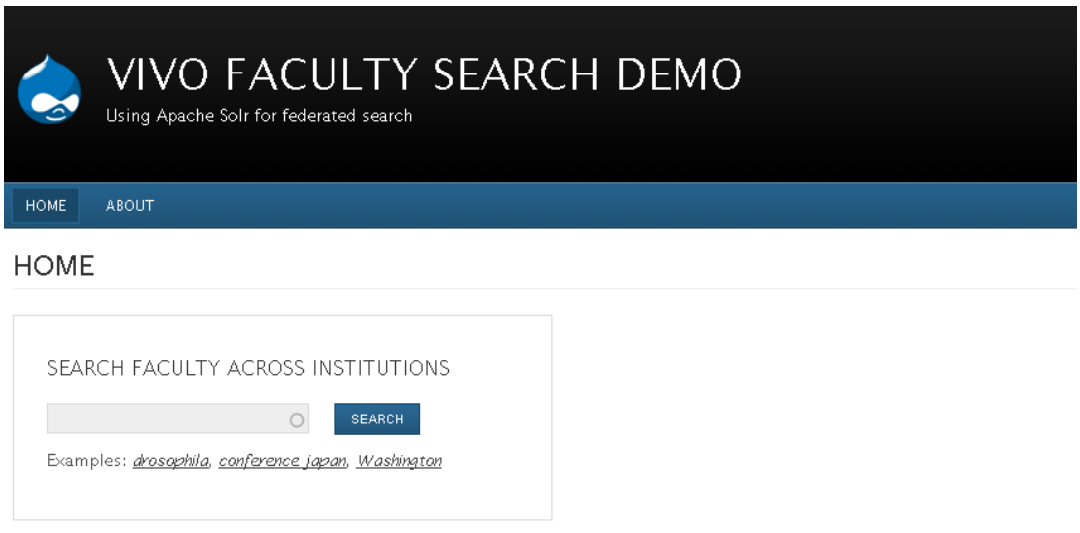


Science Map – shows where a person, department, or university publishes most in the world of science. (generated using dummy data)

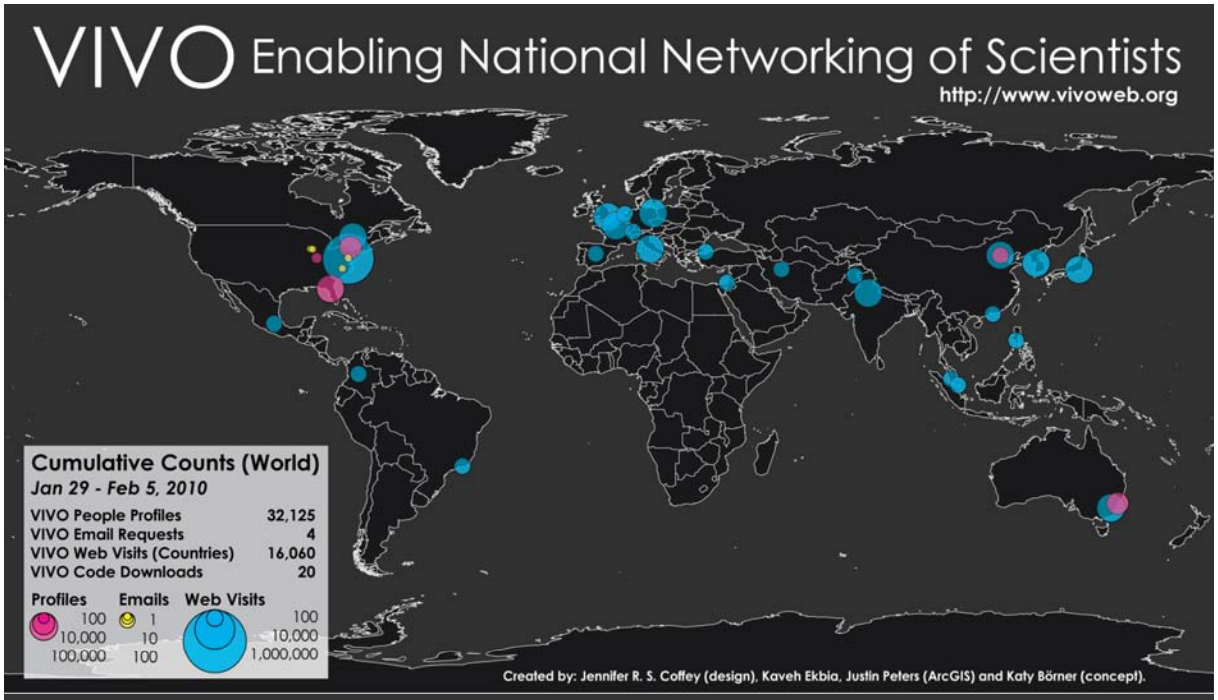
VIVO National Level Visualizations



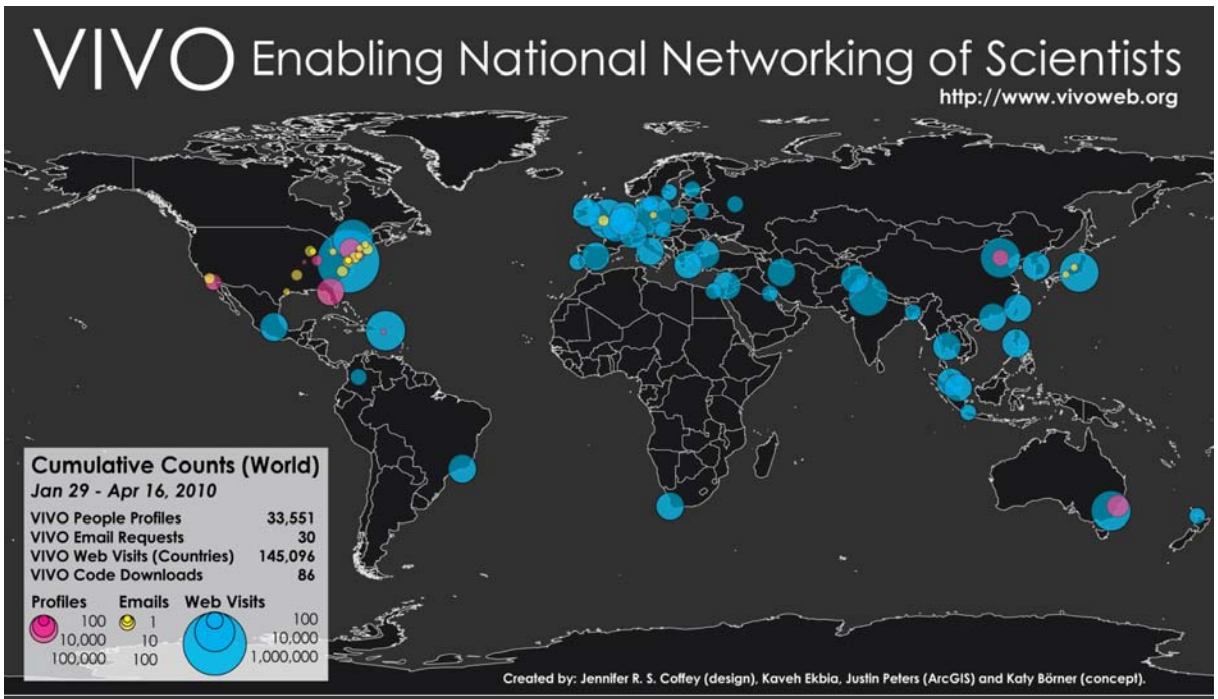
VIVO National Level Search



http://milesworthington.com/vivosearch

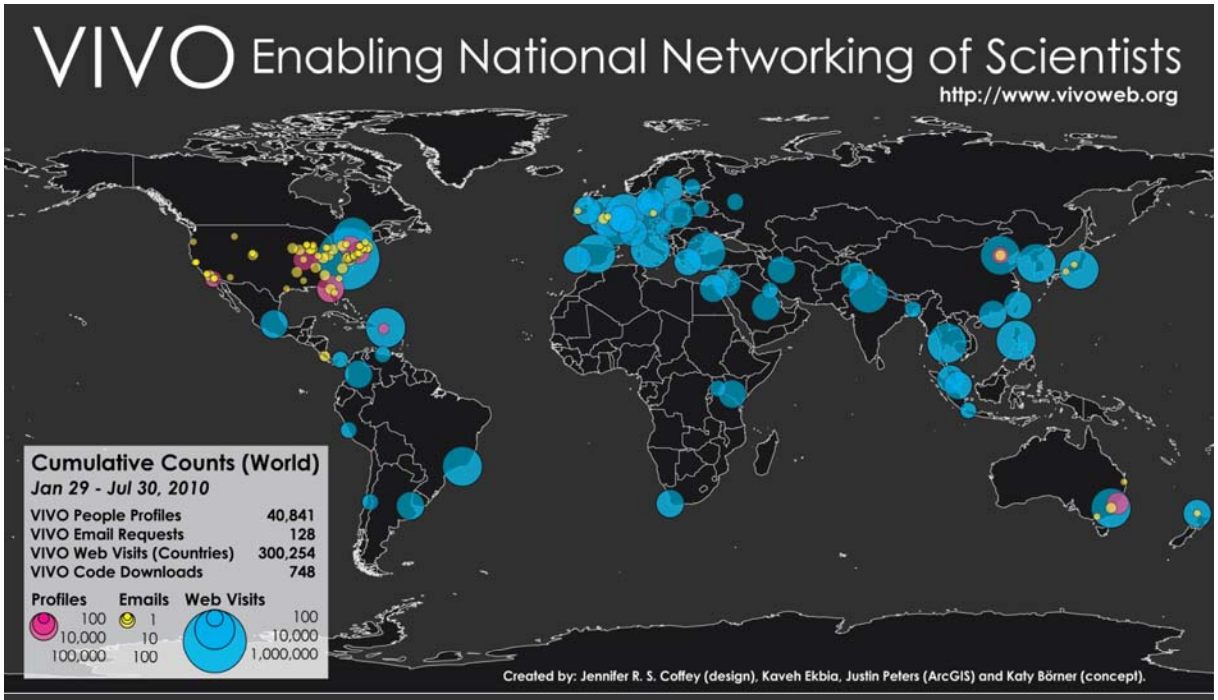


Science is global. World view of VIVO activity.
Web site visits are aggregated at the country level.



Shown are the
- Number of people profiles in the 7 different VIVO installation sites plus CAS and U Melbourne.
- Email contacts by data and service providers as well as institutions interested to adopt VIVO.
- The number of visitors on http://vivoweb.org

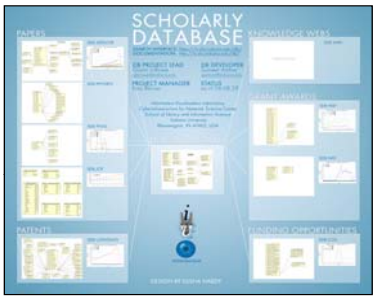
Circles are area size coded using a logarithmic scale.



VIVO 1.0 source code was publicly released on April 14, 2010
87 downloads by June 11, 2010.
The more institutions adopt VIVO, the more high quality data will be available to understand, navigate, manage, utilize, and communicate progress in science and technology.

29

Computational Scientometrics Cyberinfrastructures



Scholarly Database: 23 million scholarly records
http://sdb.slis.indiana.edu

James S. McDonnell Foundation

VIVO Research Networking
http://vivoweb.org

Information Visualization Cyberinfrastructure
http://iv.slis.indiana.edu

Network Workbench Tool & Community Wiki
http://nwb.slis.indiana.edu

Science of Science (Sci^2) Tool and CI Portal
http://sci.slis.indiana.edu

Epidemics Cyberinfrastructure
http://epic.slis.indiana.edu/

OSGi Alliance

CIShell cyberinfrastructure shell

30

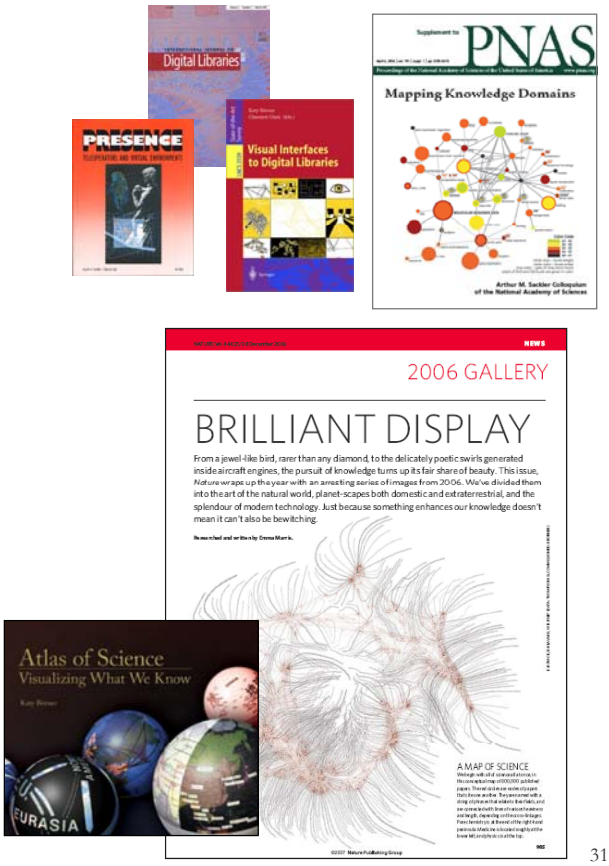
Computational Scientometrics References

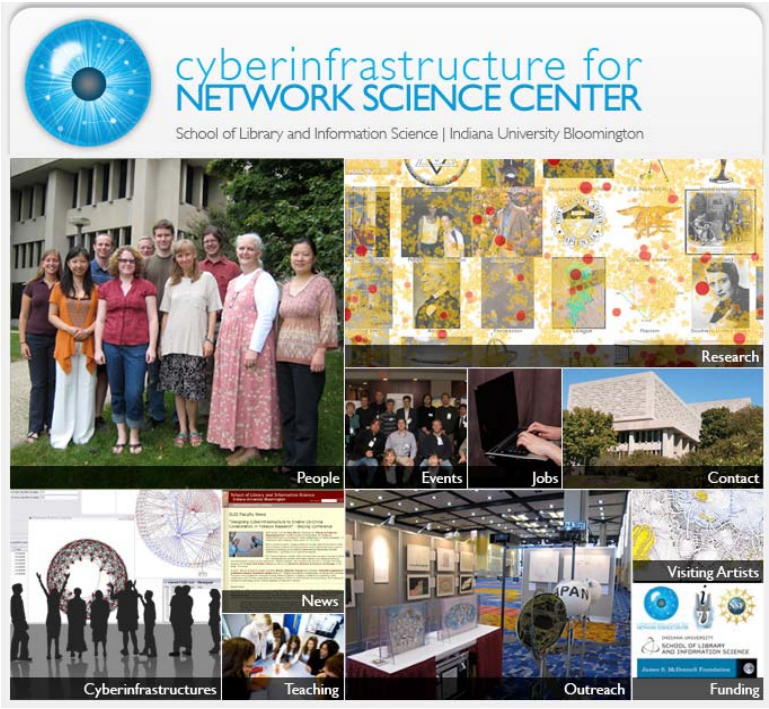
Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains.** In Blaise Cronin (Ed.), *ARIST*, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, Volume 37, Chapter 5, pp. 179-255. http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf

Shiffrin, Richard M. and Börner, Katy (Eds.) (2004). **Mapping Knowledge Domains**. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl_1). http://www.pnas.org/content/vol101/suppl_1/

Börner, Katy, Sanyal, Soma and Vespignani, Alessandro (2007). **Network Science.** In Blaise Cronin (Ed.), *ARIST*, Information Today, Inc./American Society for Information Science and Technology, Medford, NJ, Volume 41, Chapter 12, pp. 537-607. http://ivl.slis.indiana.edu/km/pub/2007-borner-arist.pdf

Börner, Katy (2010) Atlas of Science. MIT Press. http://scimaps.org/atlas





All papers, maps, cyberinfrastructures, talks, press are linked from http://cns.slis.indiana.edu