

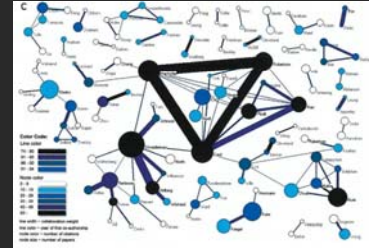
# Mapping Scientific Networks

**Dr. Katy Börner**

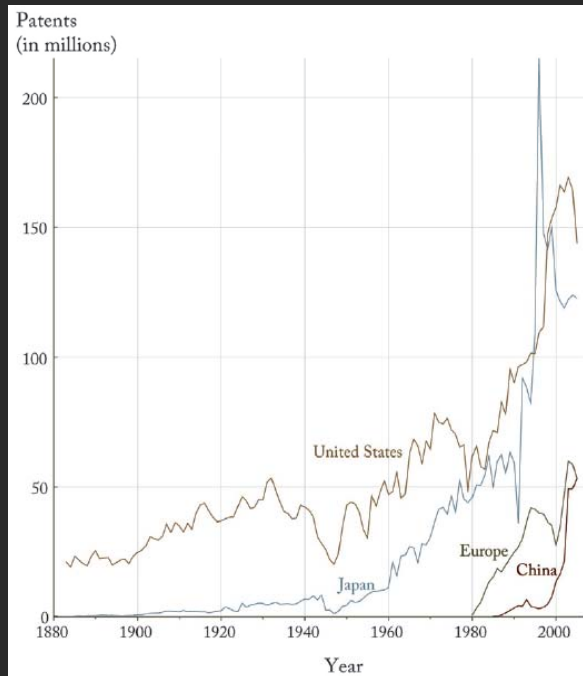
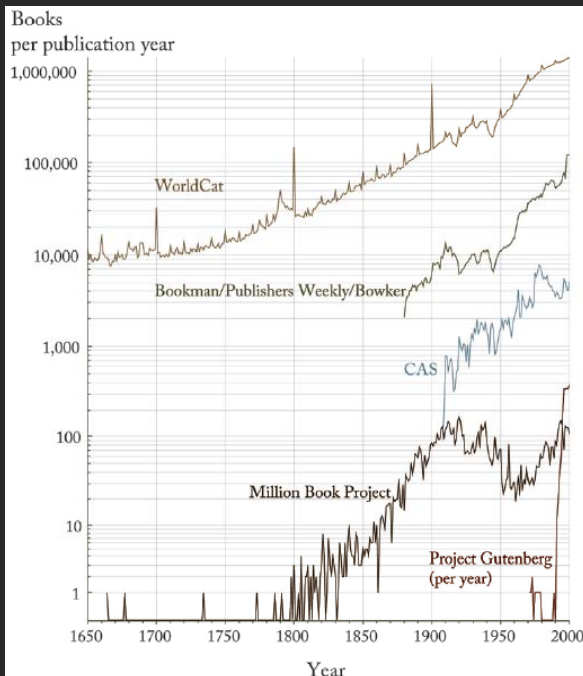
Cyberinfrastructure for Network Science Center, Director  
 Information Visualization Laboratory, Director  
 School of Library and Information Science  
 Indiana University, Bloomington, IN  
[katy@indiana.edu](mailto:katy@indiana.edu)



With special thanks to the members at the Cyberinfrastructure for Network Science Center, Mapping Science exhibit map makers and advisory board members, and the VIVO team.

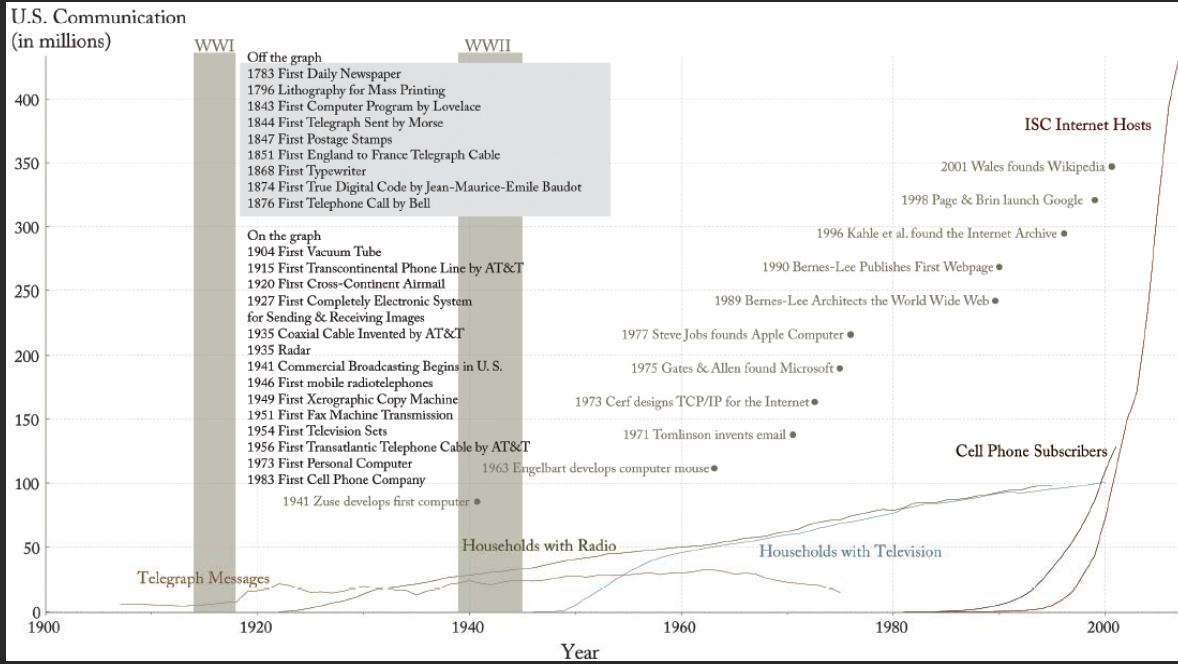


*First Annual National VIVO Conference  
 New York Hall of Science  
 August 13, 2010*

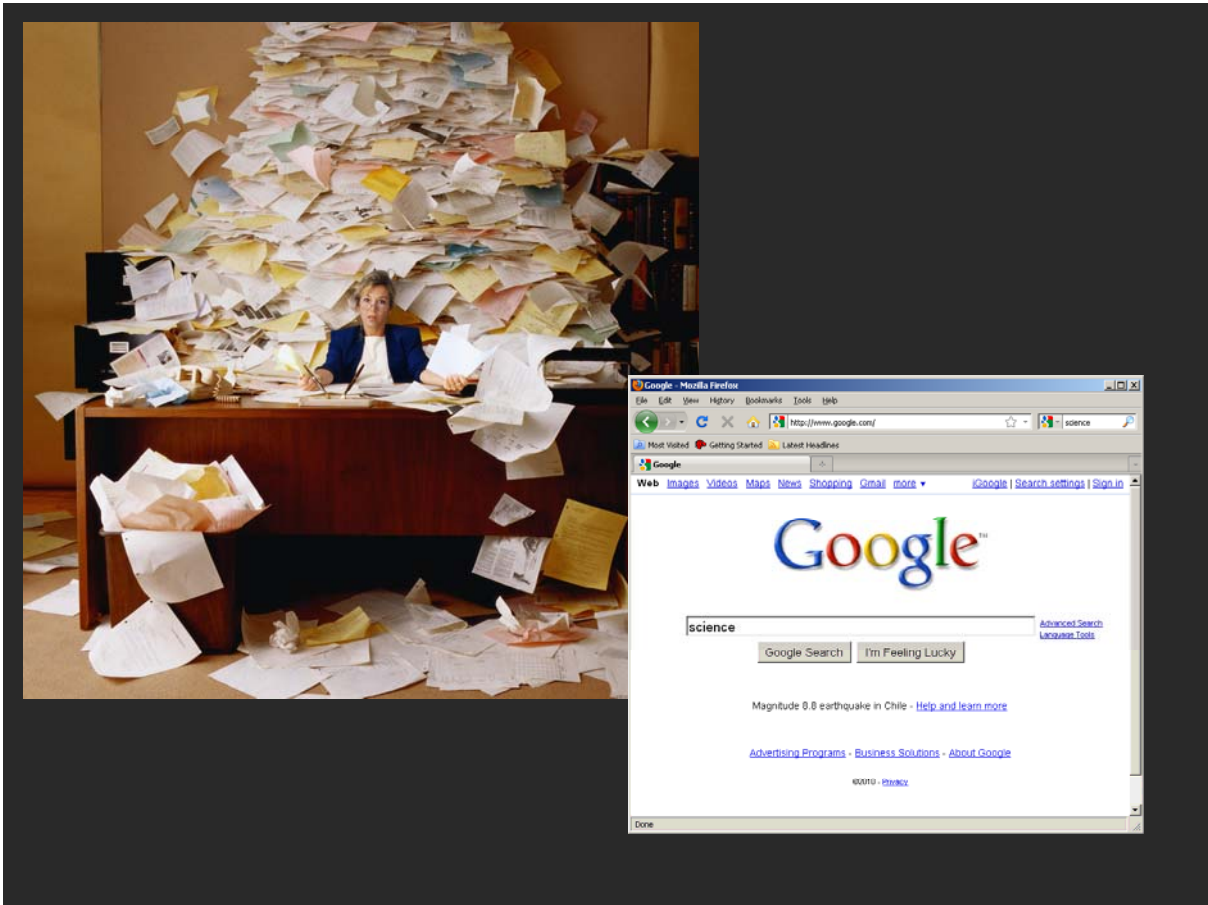


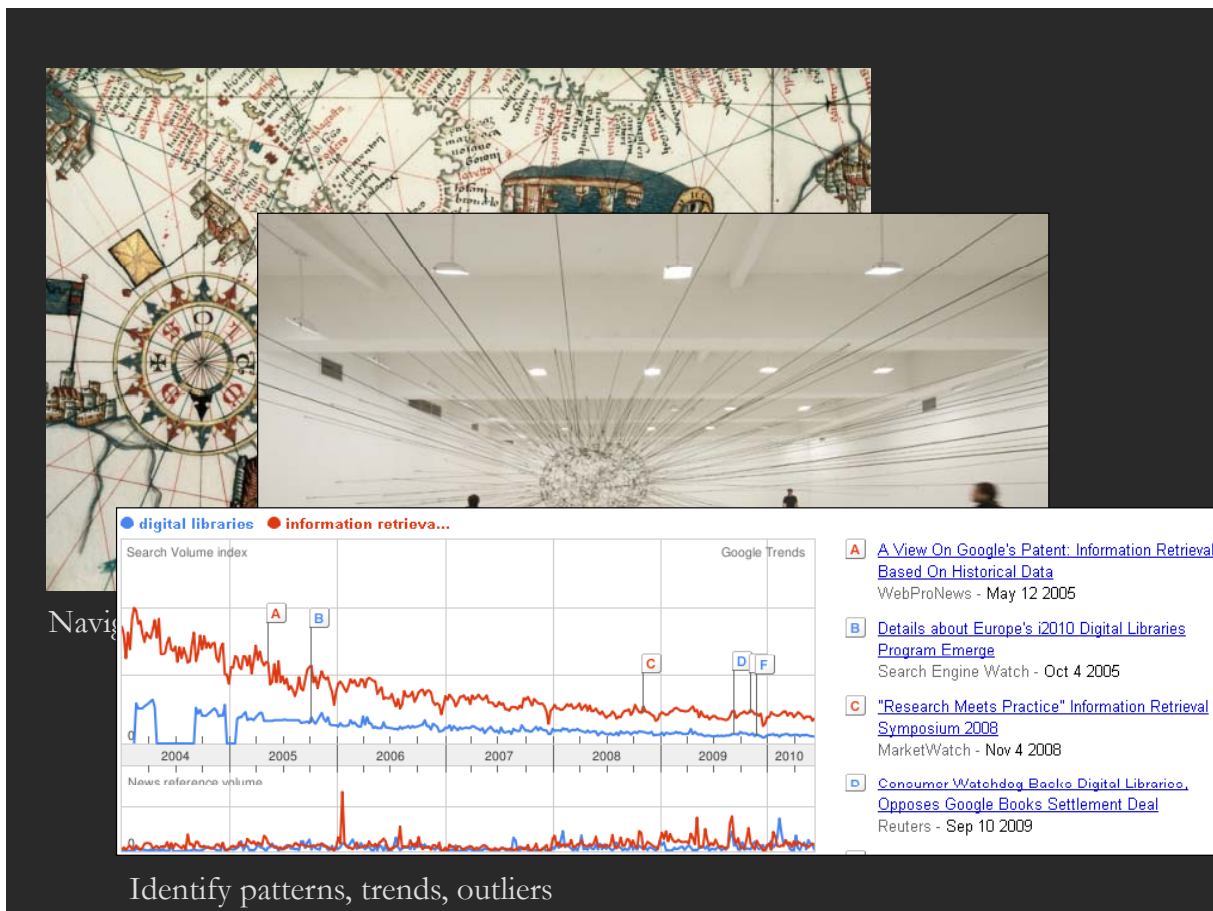
Börner, Katy (2010) *Atlas of Science*. MIT Press. <http://scimaps.org/atlas>





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





## S&T Navigation, Management Tools that Different Stakeholders Want

### Funding Agencies

- Need to monitor (long-term) money flow and research developments, identify areas for future development, stimulate new research areas, evaluate funding strategies for different programs, decide on project durations, funding patterns.

### Scholars

- Want easy access to research results, relevant funding programs and their success rates, potential collaborators, competitors, related projects/publications (**research push**).

### Industry

- Is interested in fast and easy access to major results, experts, etc. Influences the direction of research by entering information on needed technologies (**industry-pull**).

### Advantages for Publishers

- Need easy to use interfaces to massive amounts of interlinked data. Need to communicate data provenance, quality, and context.

### Society

- Needs easy access to scientific knowledge and expertise.

Some Tools That Scholars Use  
to access data, knowledge, expertise



## The Changing Scientific Landscape

**Star Scientist -> Research Teams:** In former times, science was driven by key scientists.

Today, science is driven by effectively collaborating co-author teams often comprising expertise from multiple disciplines and several geospatial locations (Börner, Dall'Asta, Ke, & Vespignani, 2005; Shneiderman, 2008).

**Users -> Contributors:** Web 2.0 technologies empower anybody to contribute to Wikipedia or to exchange images and videos via Flickr and YouTube. WikiSpecies, WikiProfessionals, or WikiProteins combine wiki and semantic technology in support of real time community annotation of scientific datasets (Mons et al., 2008).

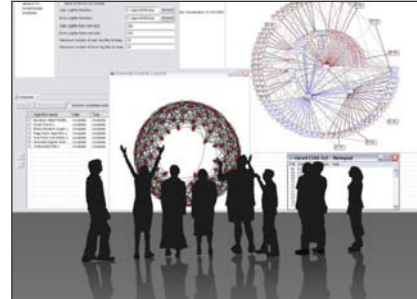
**Cross-disciplinary:** The best tools frequently borrow and synergistically combine methods and techniques from different disciplines of science and empower interdisciplinary and/or international teams of researchers, practitioners, or educators to fine-tune and interpret results collectively.

**One Specimen -> Data Streams:** Microscopes and telescopes were originally used to study one specimen at a time. Today, many researchers must make sense of massive streams of multiple types of data with different formats, dynamics, and origin.

**Static Instrument -> Evolving Cyberinfrastructure (CI):** The importance of hardware instruments that are rather static and expensive decreases relative to software infrastructures that are highly flexible and continuously evolving according to the needs of different sciences. Some of the most successful services and tools are decentralized increasing scalability and fault tolerance.



## Microscopes, Telescopes, and Macroscopes



Just as the **microscope** empowered our naked eyes to see cells, microbes, and viruses thereby advancing the progress of biology and medicine or the **telescope** opened our minds to the immensity of the cosmos and has prepared mankind for the conquest of space, **macroscopes** promise to help us cope with another infinite: the infinitely complex. Macroscopes give us a ‘vision of the whole’ and help us ‘synthesize’. They let us detect patterns, trends, outliers, and access details in the landscape of science. Instead of making things larger or smaller, macroscopes let us observe what is at once too great, too slow, or too complex for our eyes.



## Desirable Features of a Social Network Studies “Macrocope”

**Core Architecture & Plugins/Division of Labor:** Computer scientists need to design the standardized, modular, easy to maintain and extend “core architecture”. Dataset and algorithm plugins, i.e., the “filling”, are provided by those that care and know most about the data and developed the algorithms: the domain experts.

**Ease of Use:** As most plugin contributions and usage will come from non-computer scientists it must be possible to contribute, share, and use new plugins without writing one line of code. Users need guidance for constructing effective workflows from 100+ continuously changing plugins.

**Modularity:** The design of software modules with well defined functionality that can be flexibly combined helps reduce costs, makes it possible to have many contribute, and increases flexibility in tool development, augmentation, and customization.

**Standardization:** Adoption of (industry) standards speeds up development as existing code can be leveraged. It helps pool resources, supports interoperability, but also eases the migration from research code to production code and hence the transfer of research results into industry applications and products.

**Open Data and Open Code:** Lets anybody check, improve, or repurpose code and eases the replication of scientific studies.



## Type of Analysis vs. Scale of Level of Analysis

	<b>Micro/Individual</b> (1-100 records)	<b>Meso/Local</b> (101-10,000 records)	<b>Macro/Global</b> (10,000 < records)
<b>Statistical Analysis/Profiling</b>	Individual person and their expertise profiles	Larger labs, centers, universities, research states	All of NS, SA, all of sci
<b>Temporal Analysis (When)</b>	Funding portfolio of one individual	Specific bursts of PNAS	113 Years of P Research
<b>Geospatial Analysis (Where)</b>	Career trajectory of one individual	Mapping a s intellectual l	PNAS
<b>Topical Analysis (What)</b>		flows in research	VxOrd/Topic NIH funding
<b>Network Analysis (With Whom?)</b>	NSI one work of	Network	NIH's cy

13



## Macroscope for Science of Science Studies

About 5-20 algorithms are involved in one single study/workflow.

DATA EXTRACTION	UNIT OF ANALYSIS	MEASURES	LAYOUT (often one code does both similarity and ordination steps)		DISPLAY
			SIMILARITY	ORDINATION	
SEARCHES ISI INSPEC Eng Index Medline ResearchIndex Patents etc.	COMMON CHOICES Journal Document Author Term	COUNTS/FREQUENCIES Attributes (e.g. terms) Author citations Co-citations By year  THRESHOLDS By counts	SCALAR (unit by unit matrix) Direct citation Co-citation Combined linkage Co-word / co-term Co-classification  VECTOR (unit by attribute matrix) Vector space model (words/terms) Latent Semantic Analysis (words/terms) incl. Singular Value Decomp (SVD)  CORRELATION (if desired) Pearson's R on any of above	DIMENSIONALITY REDUCTION Eigenvector/ Eigenvalue solutions Factor Analysis (FA) and Principal Components Analysis (PCA) Multi-dimensional scaling (MDS) LSA, Topics Pathfinder networks (PFNet) Self-organizing maps (SOM) includes SOM, ET-maps, etc.  CLUSTER ANALYSIS  SCALAR Triangulation Force-directed placement (FDP)	INTERACTION Browse Pan Zoom Filter Query Detail on demand  ANALYSIS
BROADENING By citation By terms					

Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003) *Visualizing Knowledge Domains*. ARIST, pp. 179-255.

Domain has about 300 core researchers, 10 key data sources, 20 common tools.

Approaches/algorithms from network science, social science, political science, economics, physics, information science, webometrics, etc. are highly relevant and new ones become available every day.



## Computational Scientometrics CI



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



VIVO Research Networking  
<http://vivoweb.org>



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



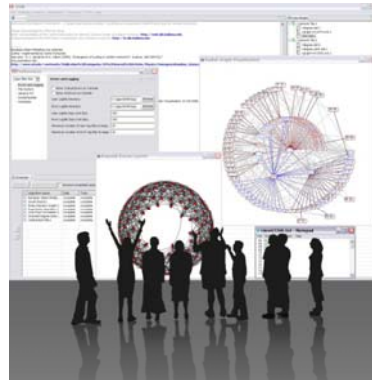
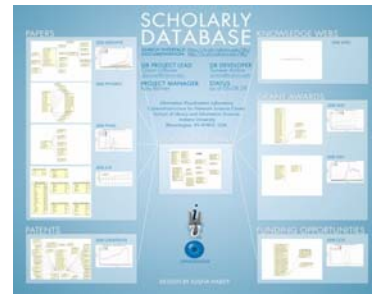
Network Workbench Tool + Community Wiki  
<http://nwb.slis.indiana.edu>



Sci<sup>2</sup> Tool and Science of Science CI Portal  
<http://sci.slis.indiana.edu>



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



15



## Macroscopic Design

### Custom Tools for Different Scientific Communities

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

**Network Workbench Tool** + Community Wiki  
<http://nwb.slis.indiana.edu>

**Science of Science (Sci<sup>2</sup>) Tool** and Portal  
<http://sci.slis.indiana.edu>

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



180+ Algorithm Plugins and Branded GUIs  
+  
Core Architecture

Open Services Gateway Initiative (OSGi) Framework.

<http://orgi.org>

Cyberinfrastructure Shell (CIShell)

<http://cishell.org>



16

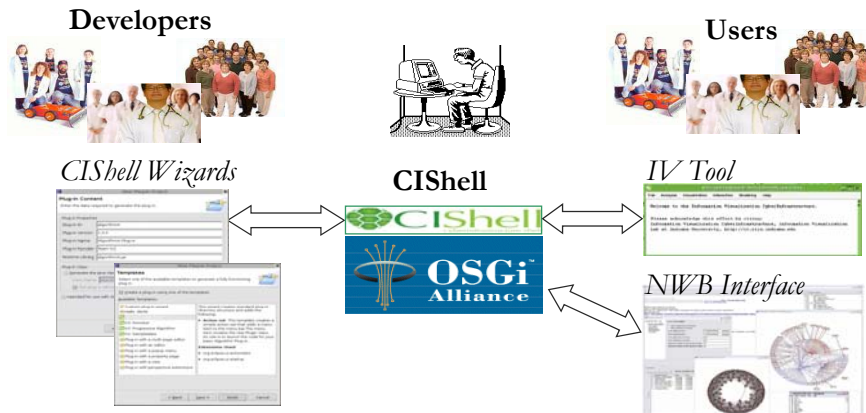




## Cyberinfrastructure Shell (CIShell)

<http://cishell.org>

- CIShell is an open source software specification for the integration and utilization of datasets, algorithms, and tools.
- It extends the Open Services Gateway Initiative (OSGi) (<http://www.osgi.org>), a standardized, component oriented, computing environment for networked services widely used in industry since 10 years.
- Specifically, CIShell provides “sockets” into which existing and new datasets, algorithms, and tools can be plugged using a wizard-driven process.



17



## CIShell – Builds on OSGi Industry Standard

CIShell is built upon the Open Services Gateway Initiative (OSGi) Framework.

### **OSGi** (<http://www.osgi.org>) is

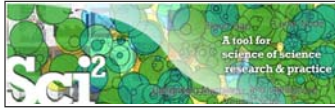
- A standardized, component oriented, computing environment for networked services.
- Successfully used in the industry from high-end servers to embedded mobile devices since 8 years.
- Alliance members include IBM (Eclipse), Sun, Intel, Oracle, Motorola, NEC and many others.
- Widely adopted in open source realm, especially since Eclipse 3.0 that uses OSGi R4 for its plugin model.

### **Advantages of Using OSGi**

- Any CIShell algorithm is a service that can be used in any OSGi-framework based system.
- Using OSGi, running CIShells/tools can be connected via RPC/RMI supporting peer-to-peer sharing of data, algorithms, and computing power.

Ideally, CIShell becomes a standard for creating OSGi Services for algorithms.

18



## Sci<sup>2</sup> Tool for Science of Science Research and Practice (<http://sci.slis.indiana.edu/sci2>)



**Forgot your password?** Katy Bömer (2010) Science of Science Research and Tools (12 Tutorials). Reporting Branch, Office of Extramural Research/Office of the Director, National Institutes of Health, Bethesda, MD.

To recover your account pas

**Not registered yet?**

[Register now](#)

**Tutorials**

- Tutorial #01: [Science of Science Research](#)
- Tutorial #02: [Network Science / Information Visualization](#)
- Tutorial #03: [CIShell Powered Tools: Network Workbench and Science of Science Tool](#)
- Tutorial #04: [Temporal Analysis—Burst Detection](#)
- Tutorial #05: [Geospatial Analysis and Mapping](#)
- Tutorial #06: [Topical Analysis & Mapping](#)
- Tutorial #07: [Tree Analysis and Visualization](#)
- Tutorial #08: [Network Analysis and Visualization](#)
- Tutorial #09: [Large Network Analysis and Visualization](#)
- Tutorial #10: [Using the Scholarly Database at IU](#)
- Tutorial #11: [YIVO National Researcher Networking](#)
- Tutorial #12: [Future Developments](#)

Scott Weingart, Hanning Guo, Bibenstine (2010) [Science of Science](#), Indiana University, I

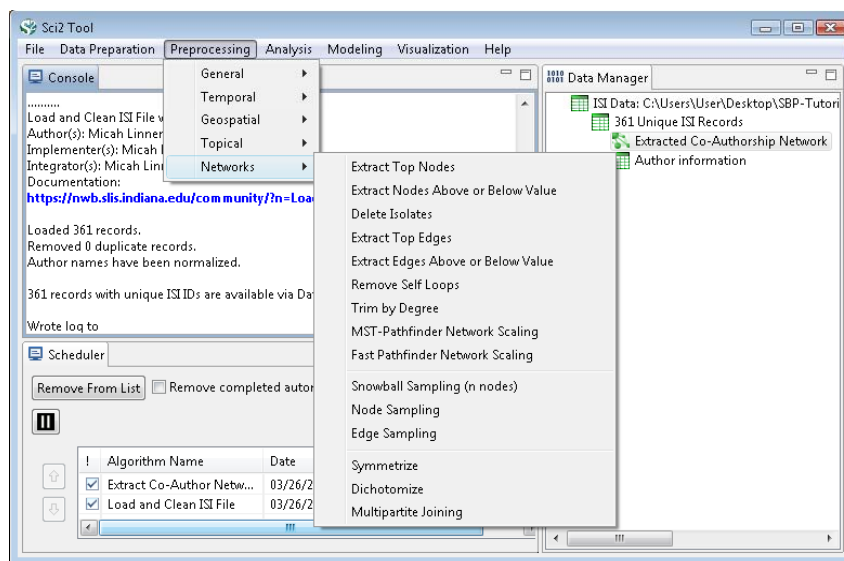
Geetha Senthil (2010). [Multidisciplinary Nature of Work With Reference to PIs and ICs Within a Portfolio](#). PA Group at NIH.

NIH Office of Extramural Research and Katy Bömer (2010) [Network Visualizations Using SPIRES Data and the Sci² Tool](#). Office of Extramural Research at NIH.

19



## Sci<sup>2</sup> Tool for Science of Science Research and Practice (<http://sci.slis.indiana.edu/sci2>)

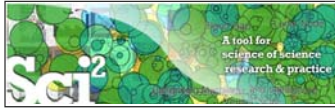


### Acknowledgments

This work is supported in part by the Cyberinfrastructure for Network Science center and the School of Library and Information Science at Indiana University, the National Science Foundation under Grant No. SBE-0738111 and IIS-0513650, and the James S. McDonnell Foundation.



20



## Sci<sup>2</sup> Tool for Science of Science

Research and Practice (<http://sci.slis.indiana.edu/sci2>)

### Supported Input file formats:

- GraphML (\*.xml or \*.graphml)
- XGMML (\*.xml)
- Pajek .NET (\*.net) & Pajek .Matrix (\*.mat)
- NWB (\*.nwb)
- TreeML (\*.xml)
- Edge list (\*.edge)
- **CSV (\*.csv)**
- **ISI (\*.isi)**
- **Scopus (\*.scopus)**
- **NSF (\*.nsf)**
- **Bibtex (\*.bib)**
- **Endnote (\*.enw)**

### Output file formats:

- GraphML (\*.xml or \*.graphml)
- Pajek .MAT (\*.mat)
- Pajek .NET (\*.net)
- NWB (\*.nwb)
- XGMML (\*.xml)
- CSV (\*.csv)

Ducu: <https://nwb.slis.indiana.edu/community/?n=DataFormats.HomePage>.

21



## Mapping the Field of RNAi Research (SDB Data)

(section 5.2.7)

<b>RNAi</b>	
<b>Time frame:</b>	1865-2008
<b>Region(s):</b>	Miscellaneous
<b>Topical Area(s):</b>	RNAi
<b>Analysis Type(s):</b>	Co-Author Network, Patent-Citation Network, Burst Detection

How many papers, patents, and funding awards exist on a specific topic?

Here we selected research on RNA interference (RNAi) is a system within living cells that helps to control which genes are active and how active they are.

The data for this analysis comes from a search of the Scholarly Database (SDB) (<http://sdb.slis.indiana.edu/>) for “RNAi” in “All Text” from MEDLINE, NSF, NIH and USPTO. A copy of this data is available in ‘\*yoursci2directory\*/sampledata/scientometrics/sdb/RNAi’. The default export format is .csv, which can be loaded in the Sci2 Tool directly.

22



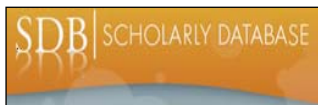
## Mapping the Field of RNAi Research (SDB Data)

(section 5.2.7)

*Email: [mwb@indiana.edu](mailto:mwb@indiana.edu)*  
*Password: mwb*

The **Scholarly Database** at Indiana University provides free access to 23,000,000 papers, patents, and grants. Since March 2009, users can also download networks, e.g., co-author, co-investigator, co-inventor, patent citation, and tables for burst analysis. For more information and to register, visit <http://sdb.slis.indiana.edu>.

23

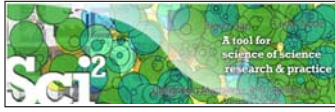


## Scholarly Database: # Records, Years Covered

Datasets available via the Scholarly Database (\* internally)

Dataset	# Records	Years Covered	Updated	Restricted Access
Medline	17,764,826	1898-2008	Yes	
PhysRev	398,005	1893-2006		Yes
PNAS	16,167	1997-2002		Yes
JCR	59,078	1974, 1979, 1984, 1989 1994-2004		Yes
USPTO	3, 875,694	1976-2008	Yes*	
NSF	174,835	1985-2004	Yes*	
NIH	1,043,804	1961-2002	Yes*	
<b>Total</b>	<b>23,167,642</b>	<b>1893-2006</b>	<b>4</b>	<b>3</b>

Aim for comprehensive time, geospatial, and topic coverage.



## Mapping the Field of RNAi Research (SDB Data)

(section 5.2.7)

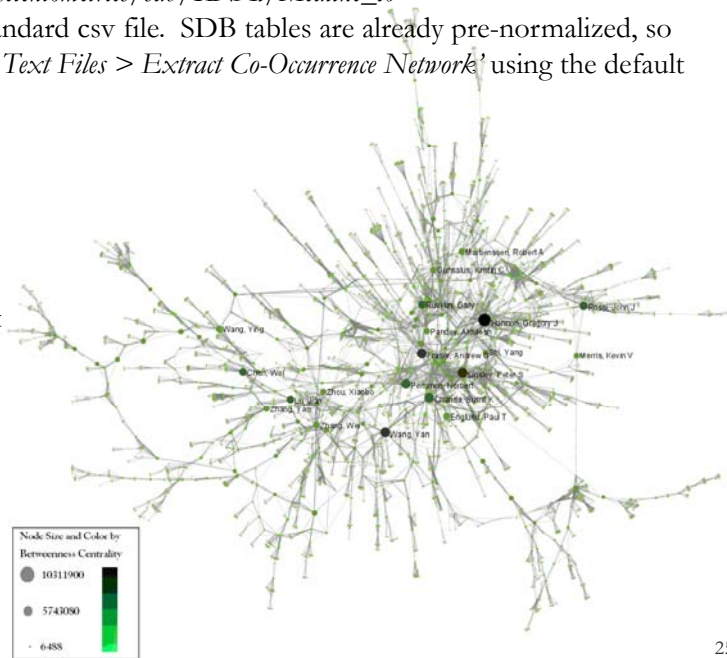
### Co-Author Network

Load `*yoursci2directory*/sampledata/scientometrics/sdb/RNAi/Medline_co-author_table_(nwb_format).csv` as a standard csv file. SDB tables are already pre-normalized, so now simply run `Data Preparation > Text Files > Extract Co-Occurrence Network` using the default parameters.

*Network Analysis Toolkit (NAT):*  
21,578 nodes with 131 isolates,  
77,739 edges.

Extract only the largest component by running `Analysis > Networks > Unweighted and Undirected > Weak Component Clustering.`

Visualize with *GUESS* using `Layout > GEM.`  
Use a custom python script to color and size the network.



25

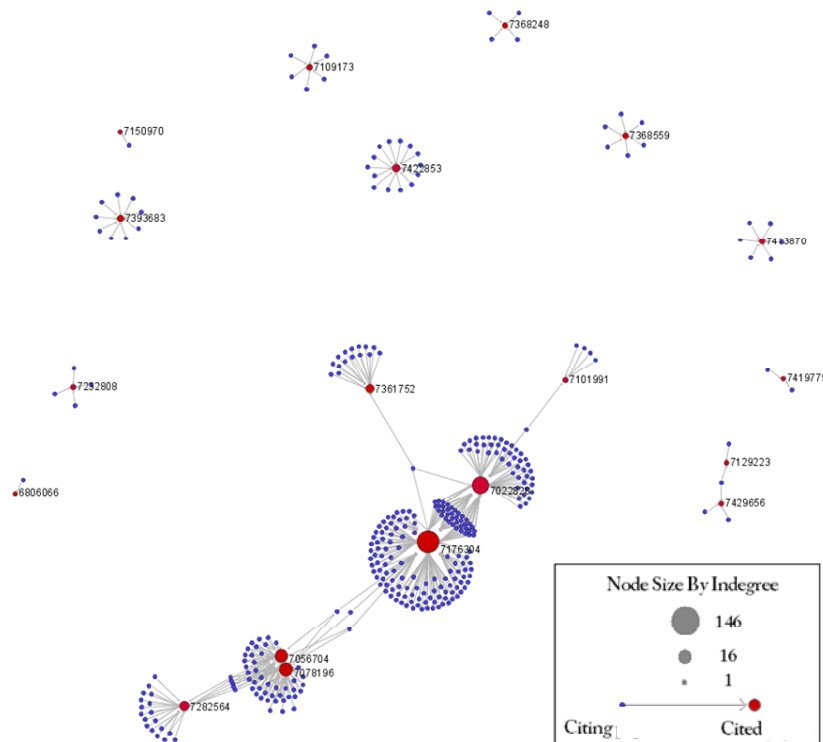


## Mapping the Field of RNAi Research (SDB Data)

(section 5.2.7)

### Patent Citation Network

To visualize the citation patterns of patents on RNAi, load `*yoursci2directory*/sampledata/scientometrics/sdb/RNAi/USPTO_citation_table_(nwb_format).csv` as a standard csv file and follow the instructions in the tutorial.



26



## Biomedical Funding Profile of NSF (NSF Data)

(section 5.2.4)

MedicalAndHealth.nsf	
<b>Time frame:</b>	2003-2010
<b>Region(s):</b>	Miscellaneous
<b>Topical Area(s):</b>	Biomedical
<b>Analysis Type(s):</b>	NSF Organization-Program Network

What organizations and programs at the National Science Foundation support projects that deal with medical and health related topics? Data was downloaded from the NSF Awards Search SIRE (<http://www.nsf.gov/awardsearch>) on Nov 23rd, 2009, using the query “medical AND health” in the title, abstract, and awards field, with “Active awards only” checked (see section 4.2.2.1 [NSF Award Search](#) for data retrieval details).

27



## Biomedical Funding Profile of NSF (NSF Data)

(section 5.2.4)

Using NSF Awards Search:  
<http://www.nsf.gov/awardsearch>  
download relevant NSF awards that have “medical” AND “health” in title, abstract, and awards. Active awards only.

Number of awards: 283 awards  
Total awarded amount to date:  
\$152,015,288

*Retrieved on Oct 18, 2009*

NSF - Award Search - Award Information - Mozilla Firefox

http://www.nsf.gov/awardsearch/awardsearch.do?searchType=PCA

National Science Foundation  
WHERE DISCOVERIES BEGIN

HOME | FUNDING | AWARDS | DISCOVERIES | NEWS | PUBLICATIONS | STATISTICS | ABOUT | FastLane

Award Search

Send Comments | Award Search Help

Awardee Information | Program Information | Search All Free-Text | Search All Fields | Show

Hint: The text field below 'Search Award For' searches the title, abstract, and award number fields.

Search Award For: "medical" and "health"

Restrict to Title Only:

Awardee Information

Principal Investigator

First Name:

Last Name:  PI Lookup

Hint: Including CO-PI will result in slower searches.

Include CO-PI:

Organization:  Organization Lookup

State:

ZIP Code:

Country:

Hint: Historical data is from prior to 1976. This data may not be as complete as recent data.

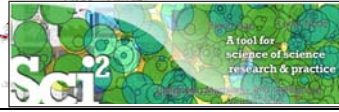
Historical Awards:

Active Awards Only:

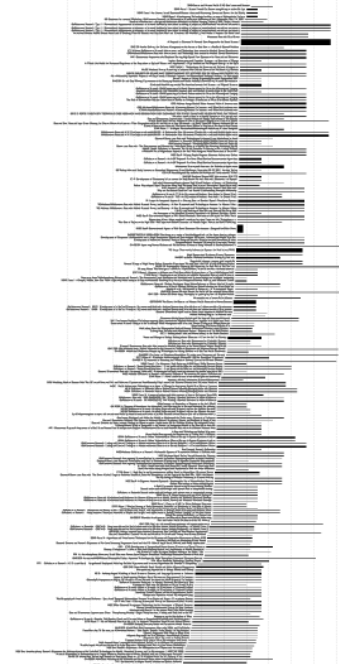
Expired Awards Only:

Search Reset

28



## Biomedical Funding Profile (section 5.2.4)



### Horizontal Bargraph

Area size equals numerical value, e.g., award amount.

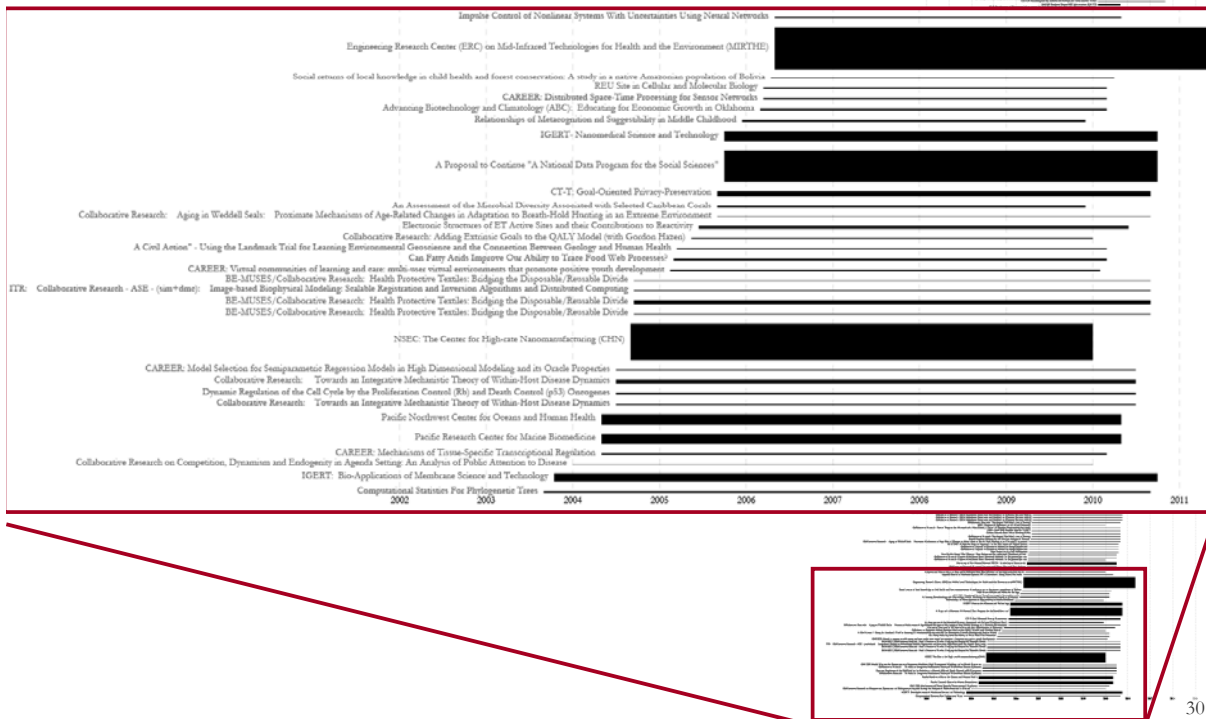


### Top-10 grants with highest \$Awarded to Date:

Title	NSF Org.	Program(s)	PI	State	Organization	\$ Awarded to Date
University of New Mexico/Harvard PREM: Leadership in Bi	DMR	PREM	MATERIALS Lopez, Gabriel	NM	University of New Mexico	2,037,500
TC: Large: Trustworthy Information Systems for Healthcare	CNS	TRUSTWORTHY	IKotz, David	NH	Dartmouth College	2,999,999
IGERT- Nanomedical Science and Technology	DGE	IGERT FULL	PF Sridhar, Srinivas	MA	Northeastern University	3,323,891
IGERT: Bio-Applications of Membrane Science and Techn	DGE	HUMAN RESOUR	Fried, Joel	OH	University of Cincinnati Main Camp	3,644,410
Pacific Research Center for Marine Biomedicine	OCE	CHEMICAL OCEA	Laws, Edward	HI	University of Hawaii	3,816,943
Pacific Northwest Center for Oceans and Human Health	OCE	CHEMICAL OCEA	Faustman, Elaine	WA	University of Washington	4,026,968
A Proposal to Continue "A National Data Program for the	SES	SCIENCE & ENG	Smith, Tom	IL	National Opinion Research Center	5,835,140
A Proposal to Continue "A National Data Program for the	SES	SCIENCE & ENG	Davis, James	IL	National Opinion Research Center	10,053,668
NSEC: The Center for High-rate Nanomanufacturing (CHN)	EEC	Studies of Policy	Busnaina, Ahmed	MA	Northeastern University	13,047,758
Engineering Research Center (ERC) on Mid-Infrared Techn	EEC	COLLABORATIVE	Gmachl, Claire	NJ	Princeton University	13,681,994



## Biomedical Funding Profile (section 5.2.4)





## Biomedical Funding Profile of NSF (NSF Data)

(section 5.2.4)

### Bimodal Network of NSF Organization to Program(s)

Extract Directed Network was selected.

Source Column: NSF Organization

Text Delimiter: |

Target Column: Program(s)

Nodes: 167

Isolated nodes: 0

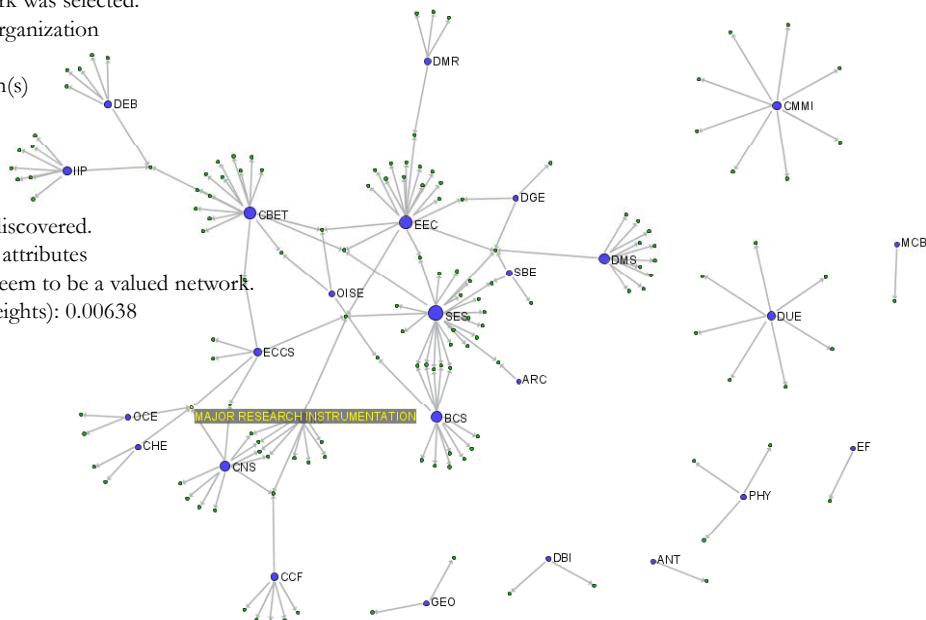
Edges: 177

No parallel edges were discovered.

Did not detect any edge attributes

This network does not seem to be a valued network.

Density (disregarding weights): 0.00638



31



## Mapping CTSA Centers (NIH RePORTER Data)

(section 5.2.3)

CTSA2005-2009.xls	
<b>Time frame:</b>	2005-2009
<b>Region(s):</b>	Miscellaneous
<b>Topical Area(s):</b>	Clinical and Translational Science
<b>Analysis Type(s):</b>	PI-Institution Network, Co-Authorship Network

A study of all NIH Clinical and Translational Science Awards (CTSA) awards and resulting publications from 2005-2009, requires advanced data acquisition and manipulation to prepare the required data. Data comes from the union of NIH RePORTER downloads (see Section 4.2.2.2 NIH RePORTER) and NIH ExPORTER data dumps (<http://projectreporter.nih.gov/exporter/>). CTSA Center grants were identified first and then matched with resulting publications using a project-specific ID. The result file is available as an Excel file in *\*yoursci2directory\*/sampledata/scientometrics/nih*. The file contains two spreadsheets, one with publication data and one with grant data. Save each spreadsheet out as *grants.csv* and *publications.csv*.

32





## Mapping CTSA Centers (NIH RePORTER Data)

(section 5.2.3)

### NIH CTSA Grants: Publication Co-Author Network

Extract Co-Occurrence Network was selected.

Input Parameters:

Text Delimiter: ;

Column Name: Authors

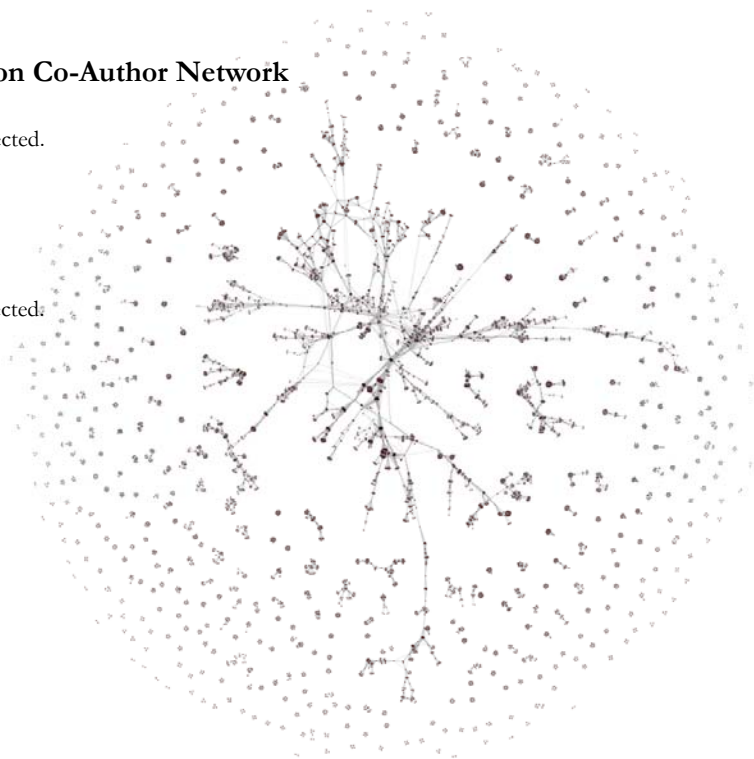
.....

Network Analysis Toolkit (NAT) was selected:

Nodes: 8680

Isolated nodes: 27

Edges: 50160



33



## Mapping CTSA Centers (NIH RePORTER Data)

(section 5.2.3)

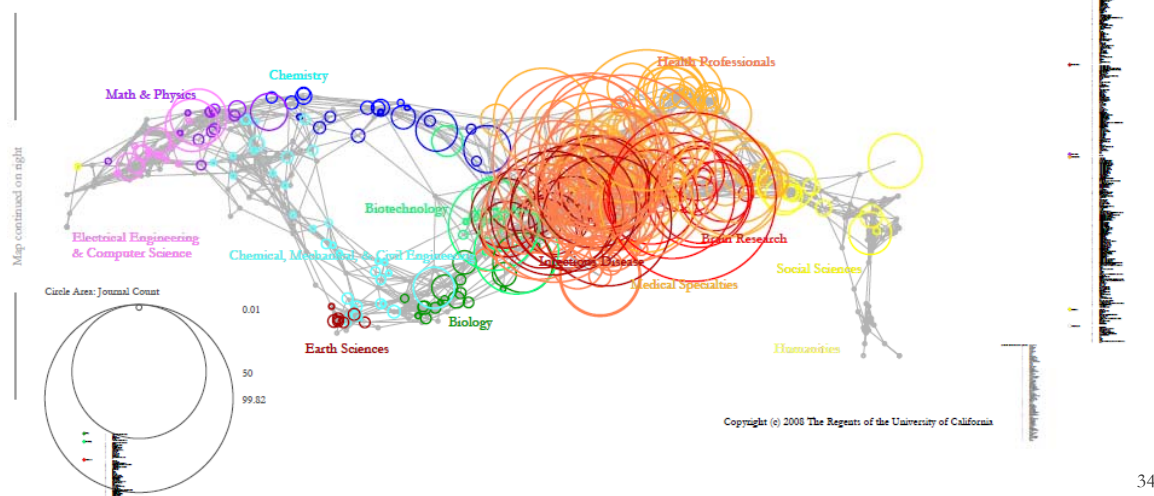
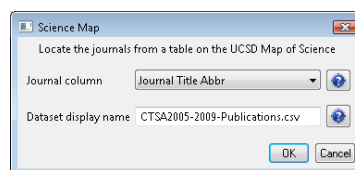
### Topic Coverage of Publications

*Visualization > Topical > Science Map via Journals'*

Science Map via Journals for CTS2005-2009-Publications.csv

2,226 journal references matched out of 2,456 found.

These 2,226 references are associated with 12 of 13 disciplines of science and 303 of 554 research specialties in the UCSD Map of Science



34

## Computational Scientometrics References

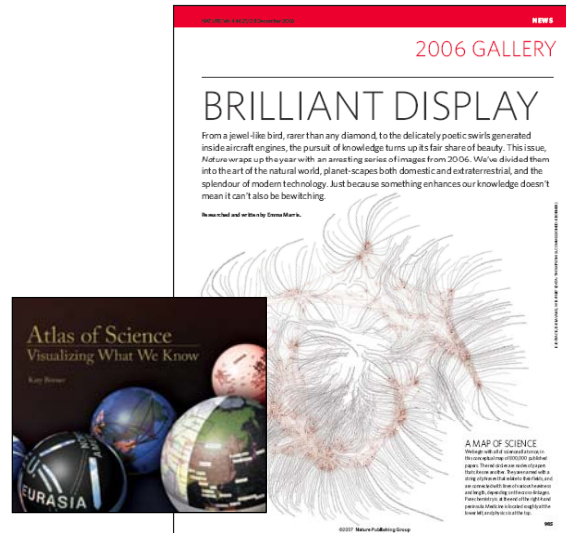
Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains**. In Blaise Cronin (Ed.), *ARIST*, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, Volume 37, Chapter 5, pp. 179-255.  
<http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf>



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/](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>

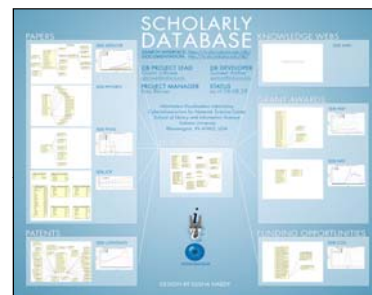
Börner, Katy (2010) *Atlas of Science*. MIT Press.  
<http://scimaps.org/atlas>



35

## Computational Scientometrics Cyberinfrastructures


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



 **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<sup>2</sup>) Tool and CI Portal**  
<http://sci.slis.indiana.edu>

 **Epidemics Cyberinfrastructure**  
<http://epic.slis.indiana.edu/>



36

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

VIVO Research & Expertise Across the University of Florida

Home People Academic Units Research Events & Seminars Index

**Conlon, Michael**  
Associate Director and Chief Operating Officer

4 publication(s) within the last 10 years

[View all VIVO publications and corresponding co-author network](#)

**Select**

Log into VIVO  
**Login**

Latest from VIVO

- [VIVO Conference: Updates](#)
- [Release 1 v. 1.1 Announcement](#)
- [VIVO on SourceForge.net: First Step to Community Development](#)
- [Reminder: early registration ends July 15!](#)
- [National VIVO Conference: Late breaking submissions announcement](#)
- [More from the VIVO blog](#)

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

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

VIVO Research & Expertise Across the University of Florida

Home People Academic Units Research Events & Seminars Index

**Conlon, Michael**  
Associate Director and Chief Operating Officer

General Statistics

- 4 publications from 1991 to 2010 (LC5VZ8)
- 12 co-author(s) from 2001 to 2010 (LC5VZ8)

Co-Author Network (LC5VZ8)

12 co-author(s)  
12 co-author(s)

**Legend**

No. of publications: 1 (blue circle), 2 (green circle), 3 (red circle)

No. of author(s) co-authored: 1 (blue line), 2 (green line), 3 (red line)

**Interact**

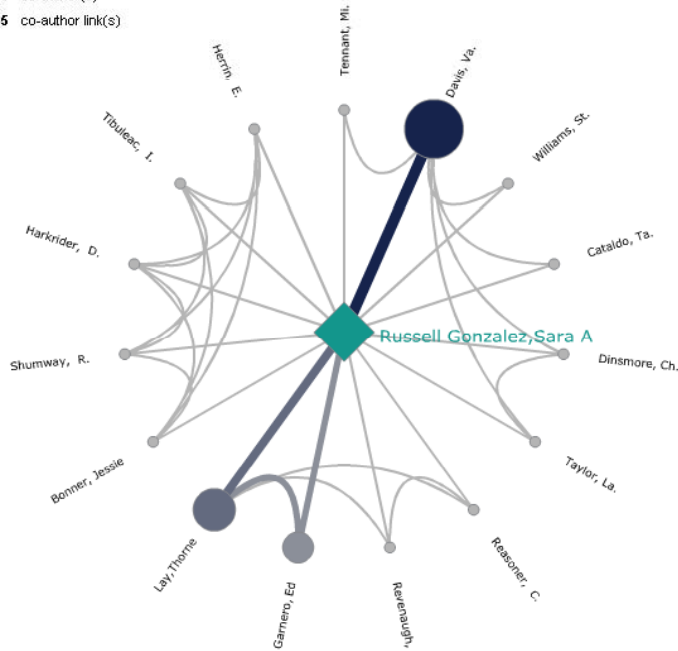
Click on an author's name to see the number of joint publications and co-author's web page. Click on a line to see details on the right.

**Tables**

Publications per year (LC5VZ8)		Co-authors (LC5VZ8)	
Year	Publications	Author	Publications with Conlon, Michael
1991	1	Watts, R. D.	2
2001	2	Shibata, H.	1
2002	1	Bachvala, M. J.	1
2004	1	Wicks, K.	1
		Chen, S. H.	1
		Shenoi, P. S. C.	1
		Chen, C. H.	1
		Pappas, C. J.	1
		Lark, A. E.	1
		Anderson, D. C.	1

**Co-Author Network** ([GraphML File](#))

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

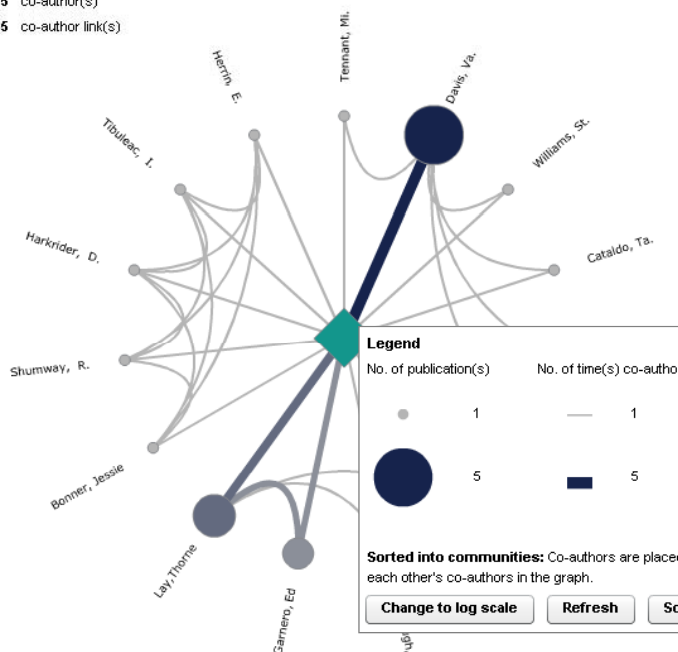


**Russell Gonzalez, Sara A**  
Physical Sciences Librarian  
[VIVO profile](#) | [Co-author network](#)

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)

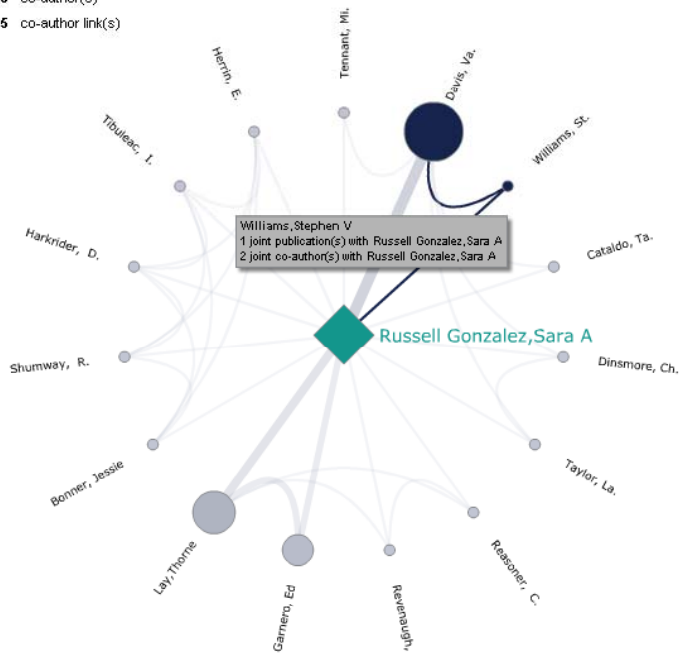


**Russell Gonzalez, Sara A**  
Physical Sciences Librarian  
[VIVO profile](#) | [Co-author network](#)

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

## Co-Author Network [\(GraphML File\)](#)

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



**Williams, Stephen V**

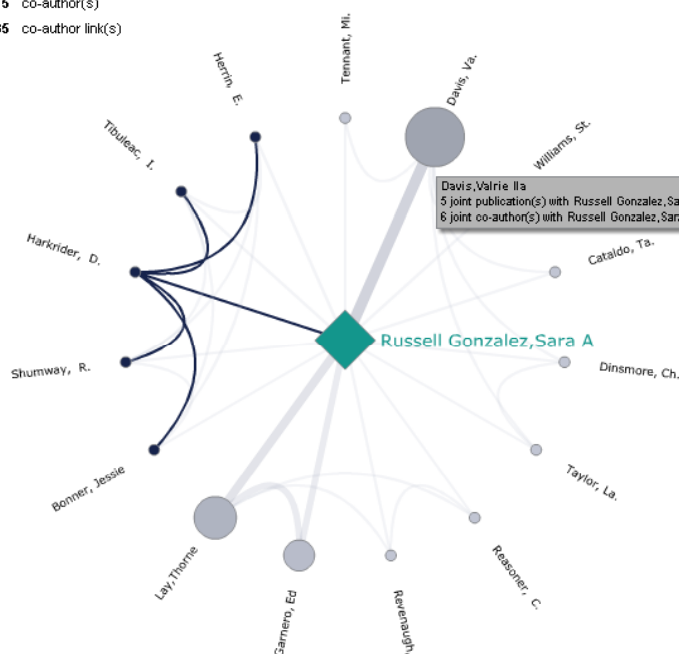
Software Engineer

[VIVO profile](#) | [Co-author network](#)

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

## Co-Author Network [\(GraphML File\)](#)

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



**Harkrider, David G.**

[VIVO profile](#) | [Co-author network](#)

1 Joint Publication(s)  
5 Joint Co-author(s)  
2003 First Publication  
2003 Last Publication

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

**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)

**Legend**

No. of publications: 0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100

No. of links (co-authored): 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100

**Interact**

Hover over any name to see the number of past publications and co-author(s) with Borner, Katy. Click on a name to see details on the right.

**Thresholding**

Only people that co-authored more than 1 paper(s) with Borner, Katy are shown.  
 15 out of 87 co-authored(s) are shown.

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

**Tables**

Publications per Year (.CSV File)

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

Co-author Publications with Borner, Katy (.CSV File)

Author	Count
Chen C.	5
Boyack K.W.	4
Mane K.K.	4
Ka W.	3
Penumarthy S.	3
Vengopran, Alessandro	2
Hein S.	2
Hardy E.	2
Holloway T.	2
Hein S. W.	2
Thakur S.	2
Feng Y.	2
Mane K.	2

## 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%2FPerson72&vis=person\\_level&render\\_mode=standalone](http://vivo-vis.slis.indiana.edu/vivo1/visualization?uri=http%3A%2F%2Fvivoweb.org%2Fontology%2Fcore%2FPerson72&vis=person_level&render_mode=standalone)

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

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)

```

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

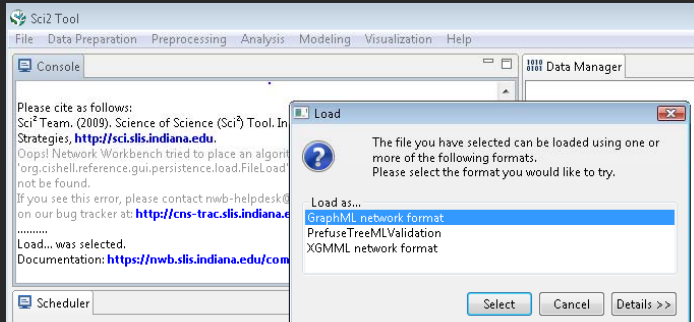
Save as Image (.PNG file)

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

Co-authors (.CSV File)

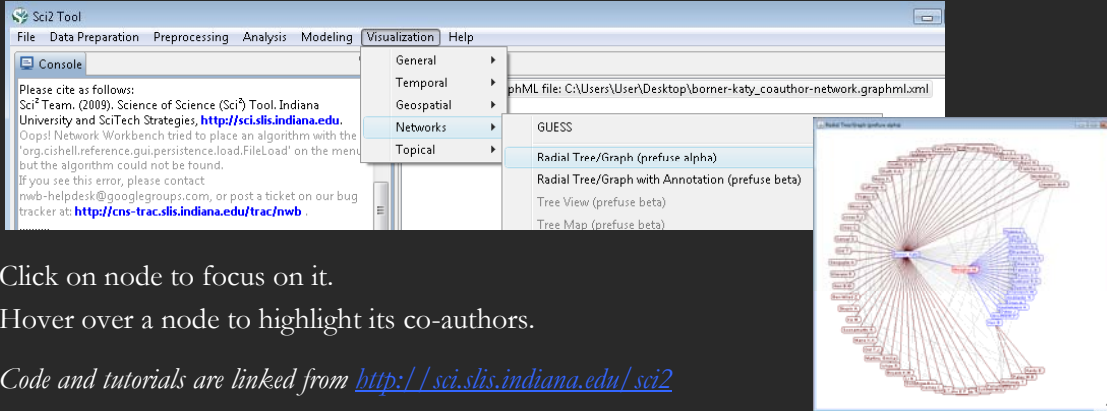
Co-Author	Count
Andrienko G.	1
Andrienko N.	1
Ben-Miled Z.	1
Blackwell A.	1
Boyack K.W.	4
Bozicevic M.	1
Brodbeck D.	1
Burkhard R.A.	1
Chen C.	5

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



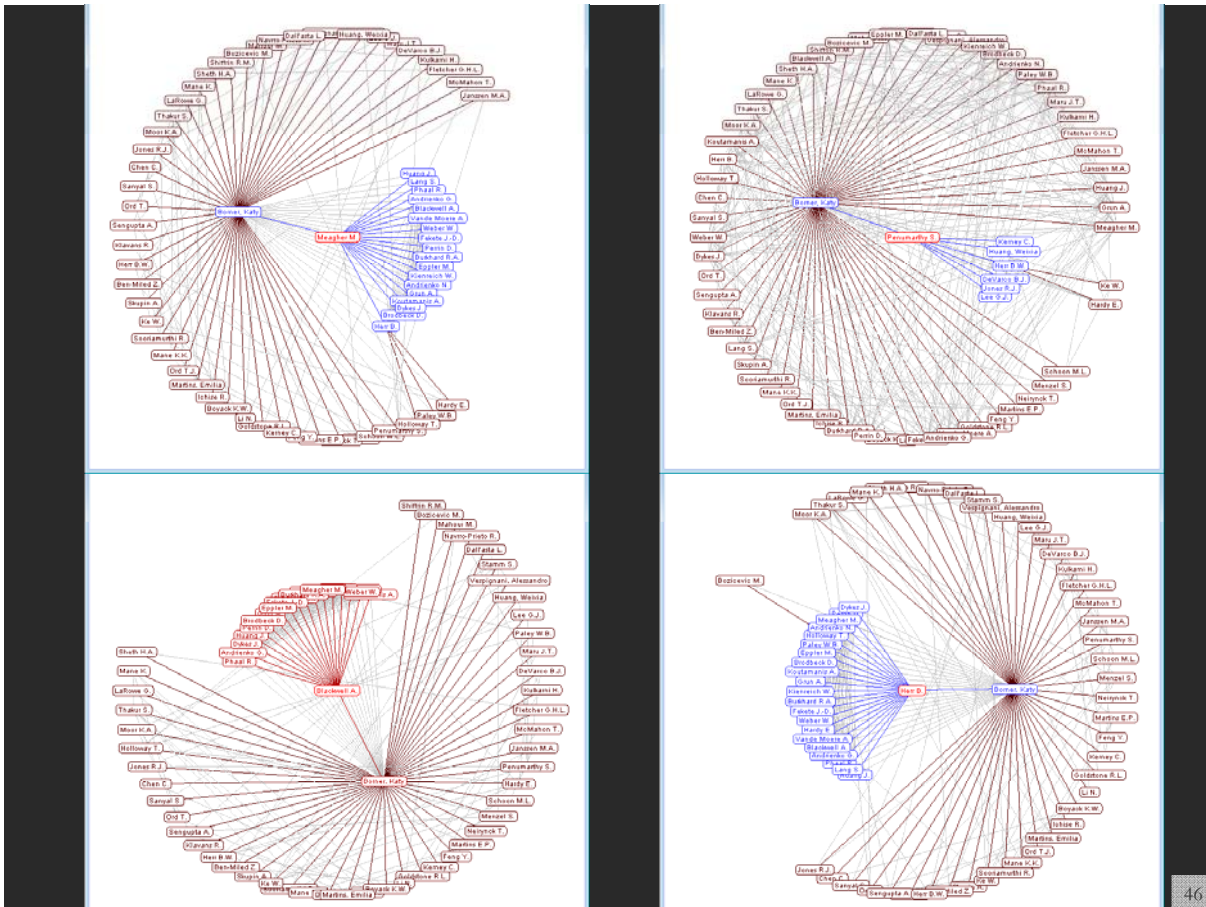
Network Analysis Toolkit  
 Nodes: 81  
 Edges: 390

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 Enabling National Networking of Scientists

<http://www.vivoweb.org>

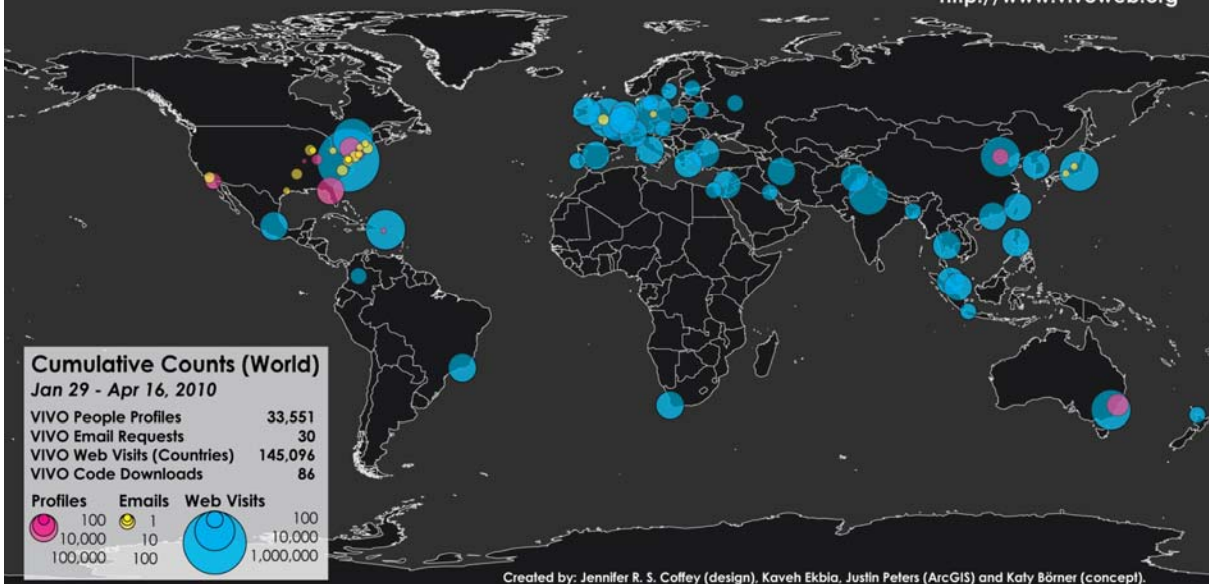


Science is global. World view of VIVO activity.  
Web site visits are aggregated at the country level.

47

# VIVO Enabling National Networking of Scientists

<http://www.vivoweb.org>



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>

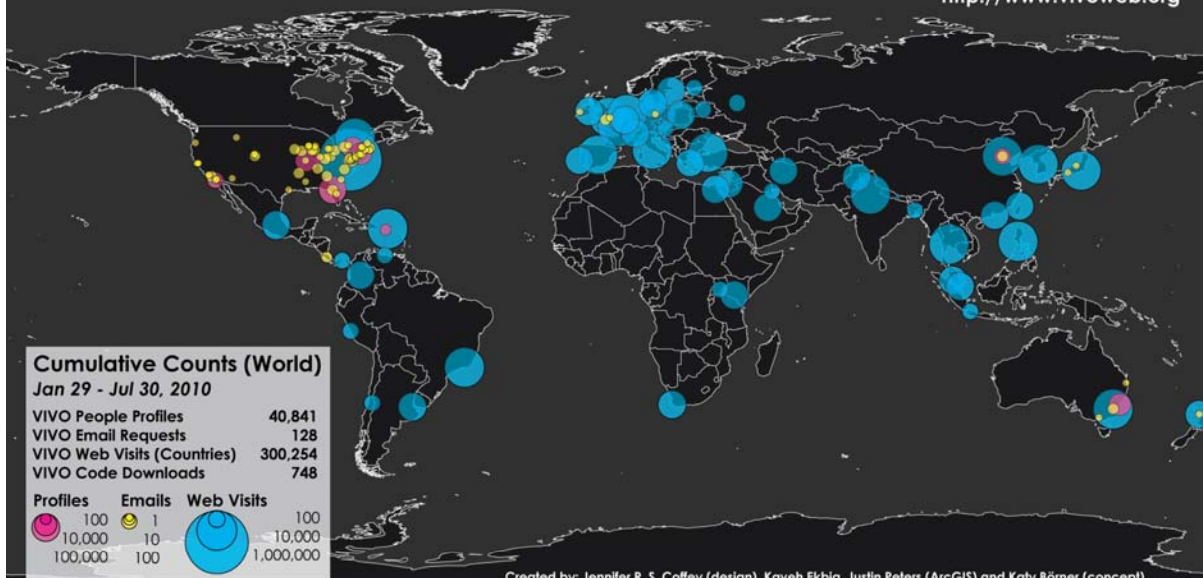
Circles are area size coded using a logarithmic scale.

48



# VIVO Enabling National Networking of Scientists

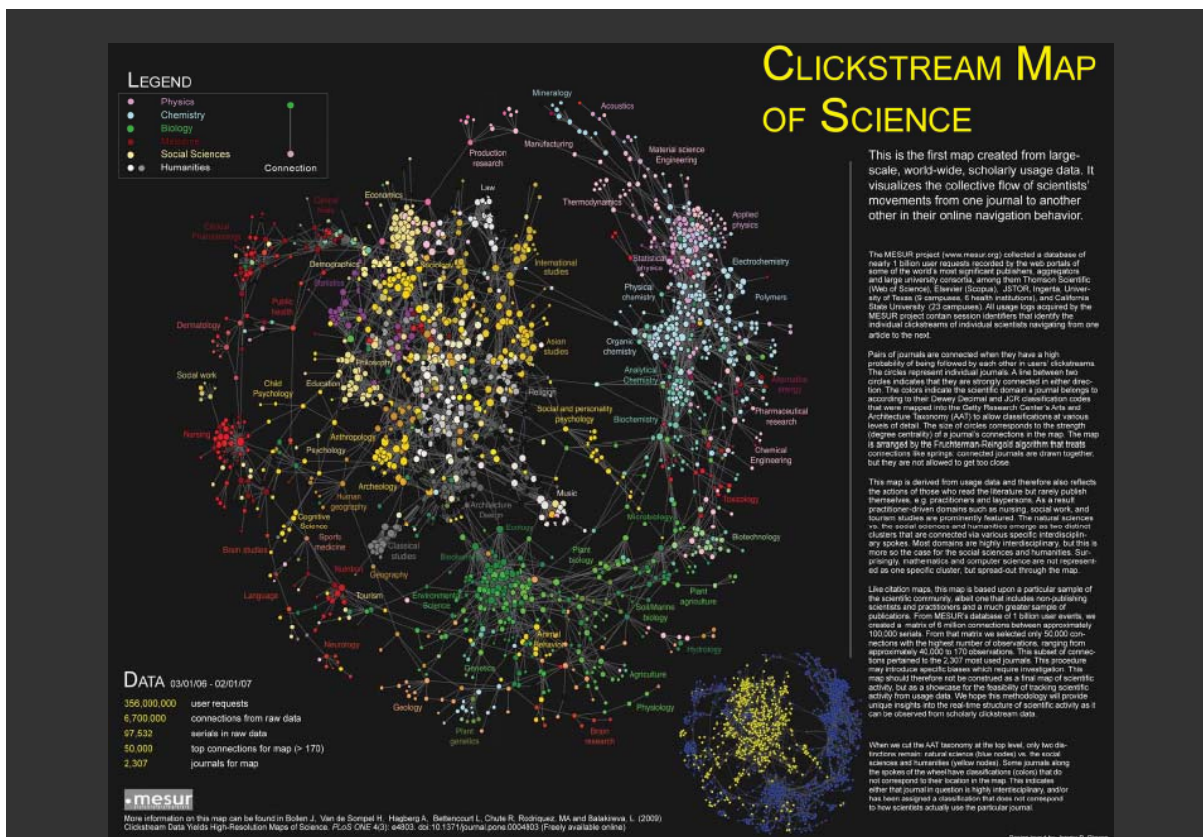
<http://www.vivoweb.org>



VIVO 1.0 source code was publicly released on April 14, 2010

87 downloads by June 11, 2010.

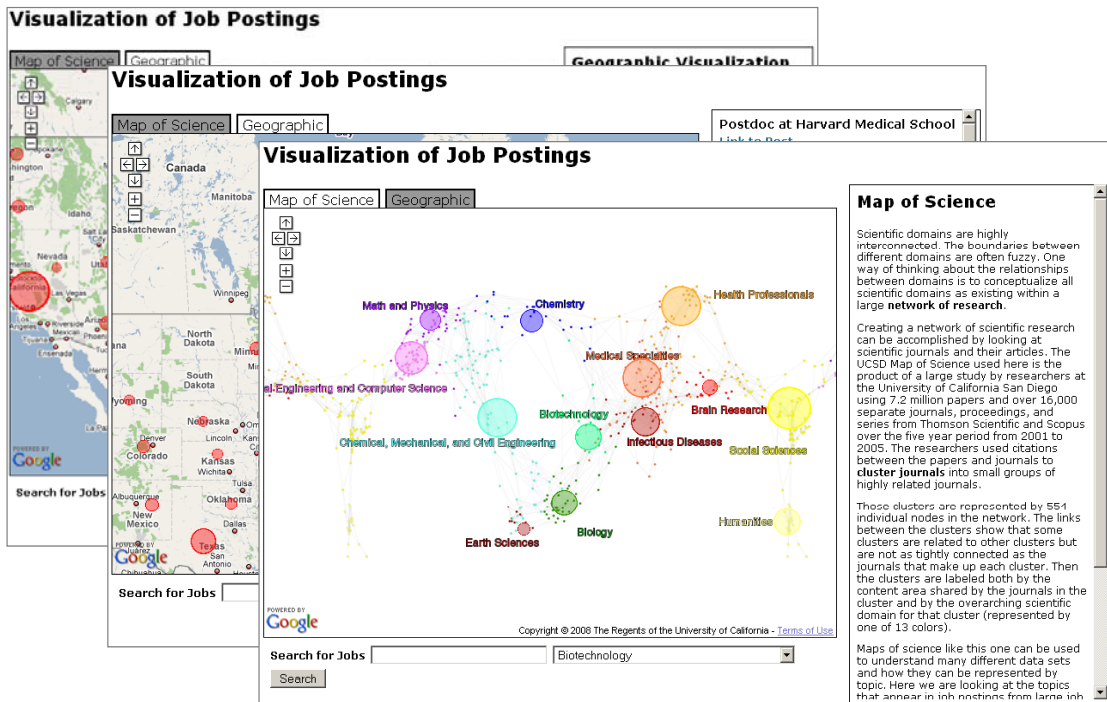
The more institutions adopt VIVO, the more high quality data will be available to understand, navigate, manage, utilize, and communicate progress in science and technology.



Bollen, Johan, Herbert Van de Sompel, Aric Hagberg, Luis M.A. Bettencourt, Ryan Chute, Marko A. Rodriguez, Lyudmila Balakireva. 2008. A Clickstream Map of Science.

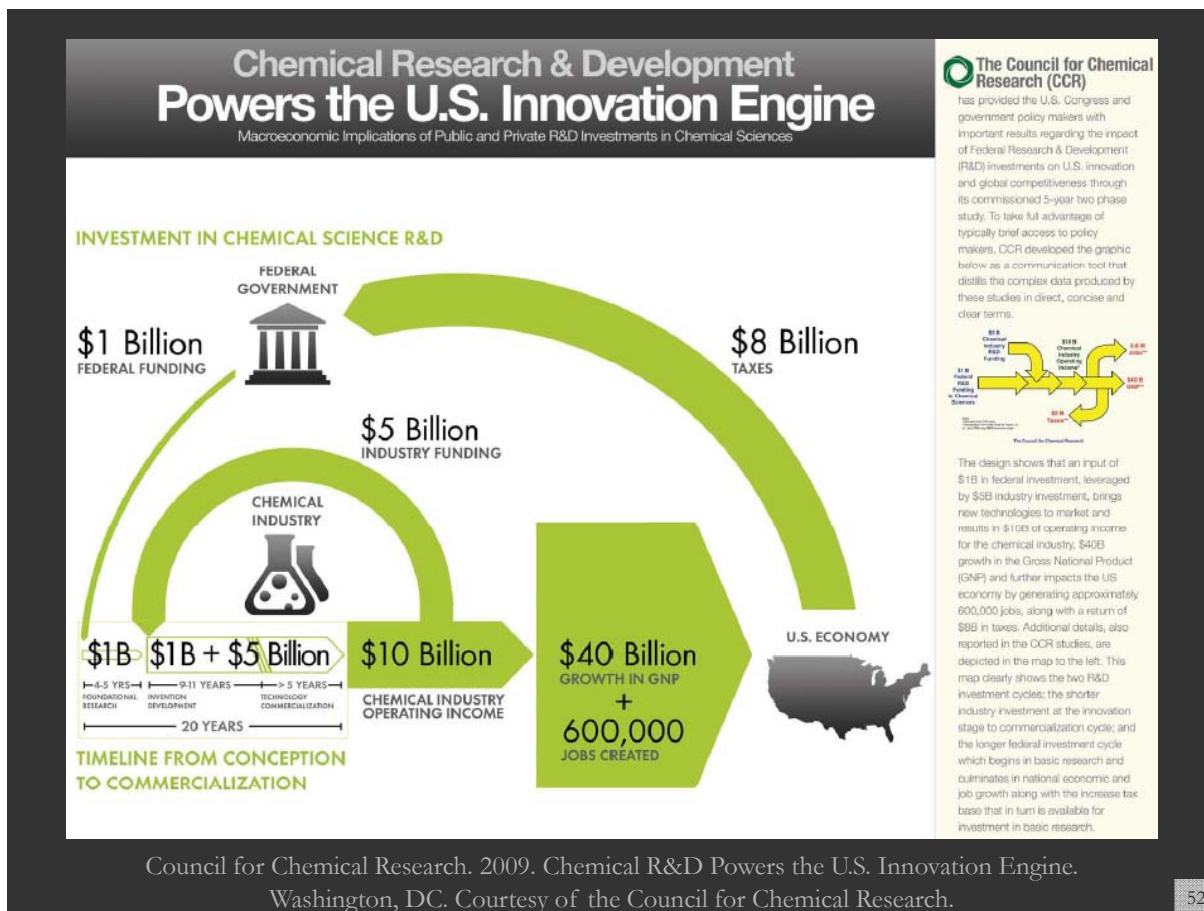
# Where Are the Academic Jobs? Interactive Exploration of Job Advertisements in Geospatial and Topical Space

Angela Zoss, Michael Connover, Katy Börner (2010)



<http://cns-nd3.slis.indiana.edu/mapjobs/geo>

51



52

# Mapping Science Exhibit – 10 Iterations in 10 years

<http://scimaps.org>

## The Power of Maps (2005)



## The Power of Reference Systems (2006)



## The Power of Forecasts (2007)



## Science Maps for Economic Decision Makers (2008)



## Science Maps for Science Policy Makers (2009)



## Science Maps for Scholars (2010)

## Science Maps as Visual Interfaces to Digital Libraries (2011)

## Science Maps for Kids (2012)

## Science Forecasts (2013)

## How to Lie with Science Maps (2014)

Exhibit has been shown in 72 venues on four continents. Currently at

- NSF, 10th Floor, 4201 Wilson Boulevard, Arlington, VA
- Marston Science Library, University of Florida, Gainesville, FL
- Center of Advanced European Studies and Research, Bonn, Germany
- Science Train, Germany.



53

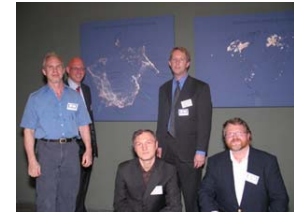


Debut of 5<sup>th</sup> Iteration of Mapping Science Exhibit at MEDIA X was on May 18, 2009 at Wallenberg Hall, Stanford University, <http://mediax.stanford.edu>, <http://scaleindependentthought.typepad.com/photos/scimaps>

54

# Illuminated Diagram Display

W. Bradford Paley, Kevin W. Boyack, Richard Kalvans, and Katy Börner (2007)  
*Mapping, Illuminating, and Interacting with Science. SIGGRAPH 2007.*



Large-scale, high resolution prints illuminated via projector or screen.

## Questions:

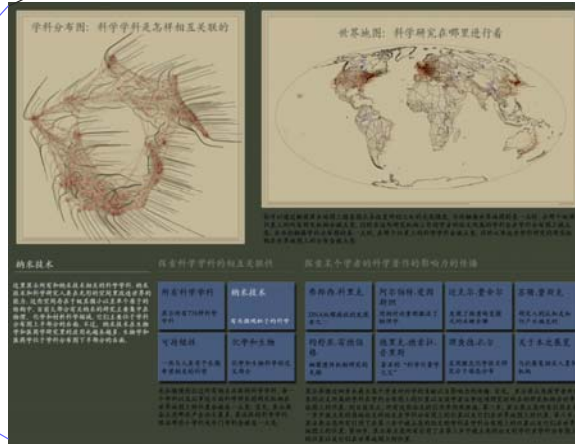
- Who is doing research on what topic and where?
- What is the ‘footprint’ of interdisciplinary research fields?
- What impact have scientists?



Interactive touch panel.

## Contributions:

- Interactive, high resolution interface to access and make sense of data about scholarly activity.



TOPIC MAP: HOW SCIENTIFIC PARADIGMS RELATE

GEOGRAPHIC MAP: WHERE SCIENCE GETS DONE

You may run your finger over each of these maps to control the lighting on the other: touching a place on the world map will light up topics studied in that place; touching a paradigm on the topic map will light up the places that study that topic.

**Nanotechnology**

This overlay shows the distribution of nanotechnology within the paradigms of science. The majority of current work in nanotechnology takes place in physics, chemistry, and materials science, at the upper right portion of the map. However, an increasing amount of nanotechnology is being applied in the biological and medical sciences, at the lower right.

<p><b>All Topics</b></p> <p>Sweep through all 276 scientific paradigms</p>	<p><b>Nanotechnology</b></p> <p>Science on the tiny scale of molecules</p>	<p><b>Francis H. C. CRICK</b></p> <p>Co-discovered DNA's double helix</p>	<p><b>Albert EINSTEIN</b></p> <p>Revitalized physics with Relativity theories</p>	<p><b>Michael E. FISHER</b></p> <p>Models critical phase transitions of matter</p>	<p><b>Susan T. FISKE</b></p> <p>Connects perception and stereotypes</p>
<p><b>Sustainability</b></p> <p>The science behind our long-term hopes</p>	<p><b>Biology &amp; Chemistry</b></p> <p>The interface between these two vital fields</p>	<p><b>Joshua LEDERBERG</b></p> <p>Pioneer in bacterial genetic mechanisms</p>	<p><b>Derek J. de Solla PRICE</b></p> <p>Known as the "Father of Scientometrics"</p>	<p><b>Richard N. ZARE</b></p> <p>Uses laser chemistry in molecular dynamics</p>	<p><b>About this display</b></p> <p>People &amp; organizations that helped create it</p>

We sweep slowly through adjoining related topics, lighting up the places in the world that study each topic. You may select a subset of the topics that deal with these three interesting subjects by touching it.

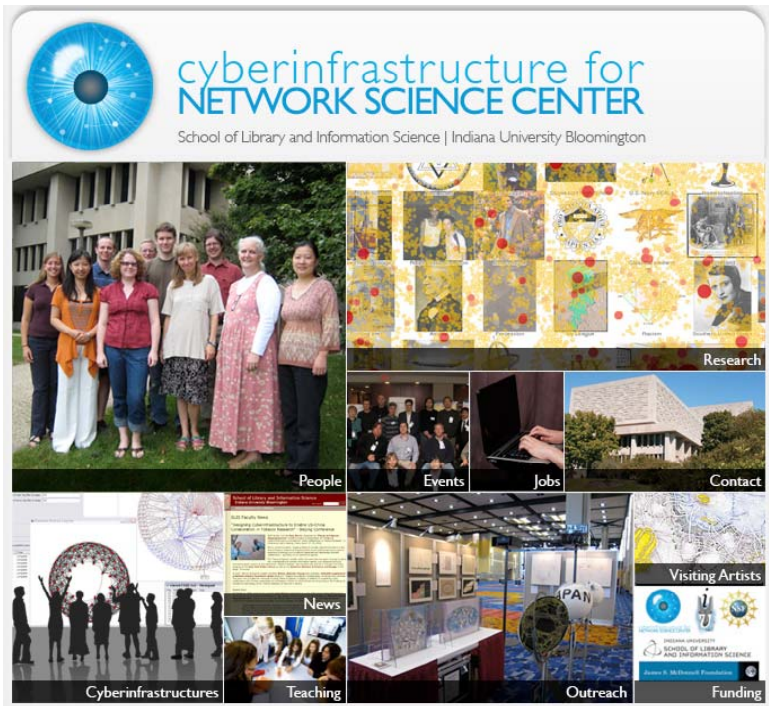
A single person's spreading influence is shown as a series of four snapshots. First, we light only topics and places relating to that person's papers—papers that are still highly cited today. The second lights everything that cites that original work. Note that this first-generation impact extends to far more topics than did the original work. The third snapshot lights science that cites the second, and the fourth lights science that cites the third.



Science Maps in “Expedition Zukunft” science train visiting 62 cities in 7 months  
12 coaches, 300 m long  
Opening was on April 23<sup>rd</sup>, 2009 by German Chancellor Merkel  
<http://www.expedition-zukunft.de>



This is the only mockup in this slide show.  
Everything else is available today.



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