

# Mining, Mapping, and Accelerating Science and Technology

Katy Börner

Royal Netherlands Academy of Arts and Sciences(KNAW), The Netherlands  
Cyberinfrastructure for Network Science Center, SLIS, IUB, USA

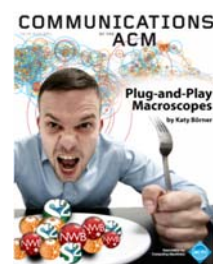
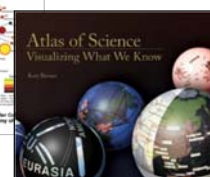
[katy@indiana.edu](mailto:katy@indiana.edu) | <http://cns.iu.edu>

KNAW, Het Trippenhuis, Amsterdam, The Netherlands  
June 6, 2012



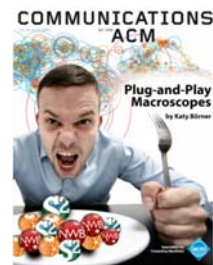
## Overview

1. **Data mining and visualization research** that aims to increase our scientific understanding of the structure and dynamics of science and technology.
2. **Novel approaches and services** that improve information access, researcher networking, and research management.
3. **Data services and plug-and-play macroscope tools** that commoditize data mining and visualization.

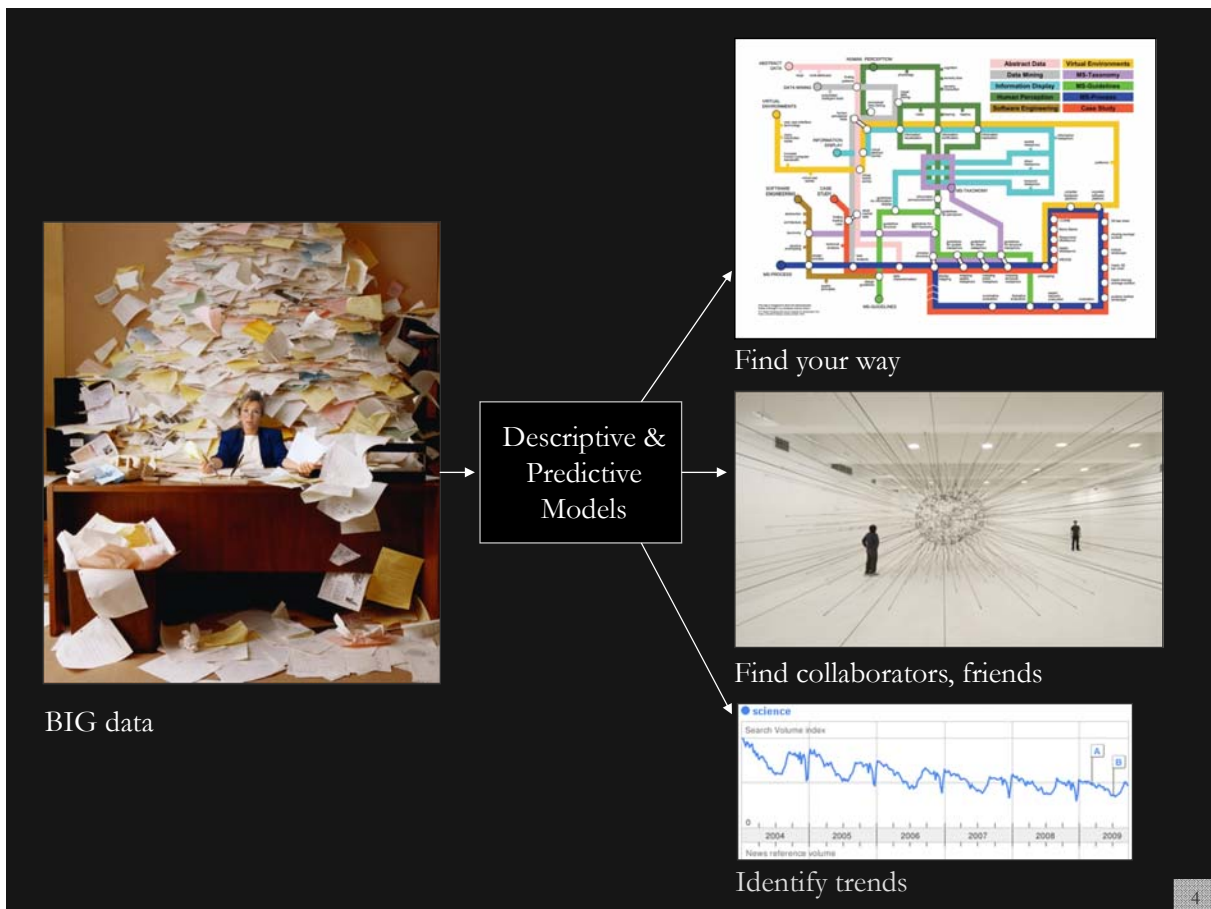


## Overview

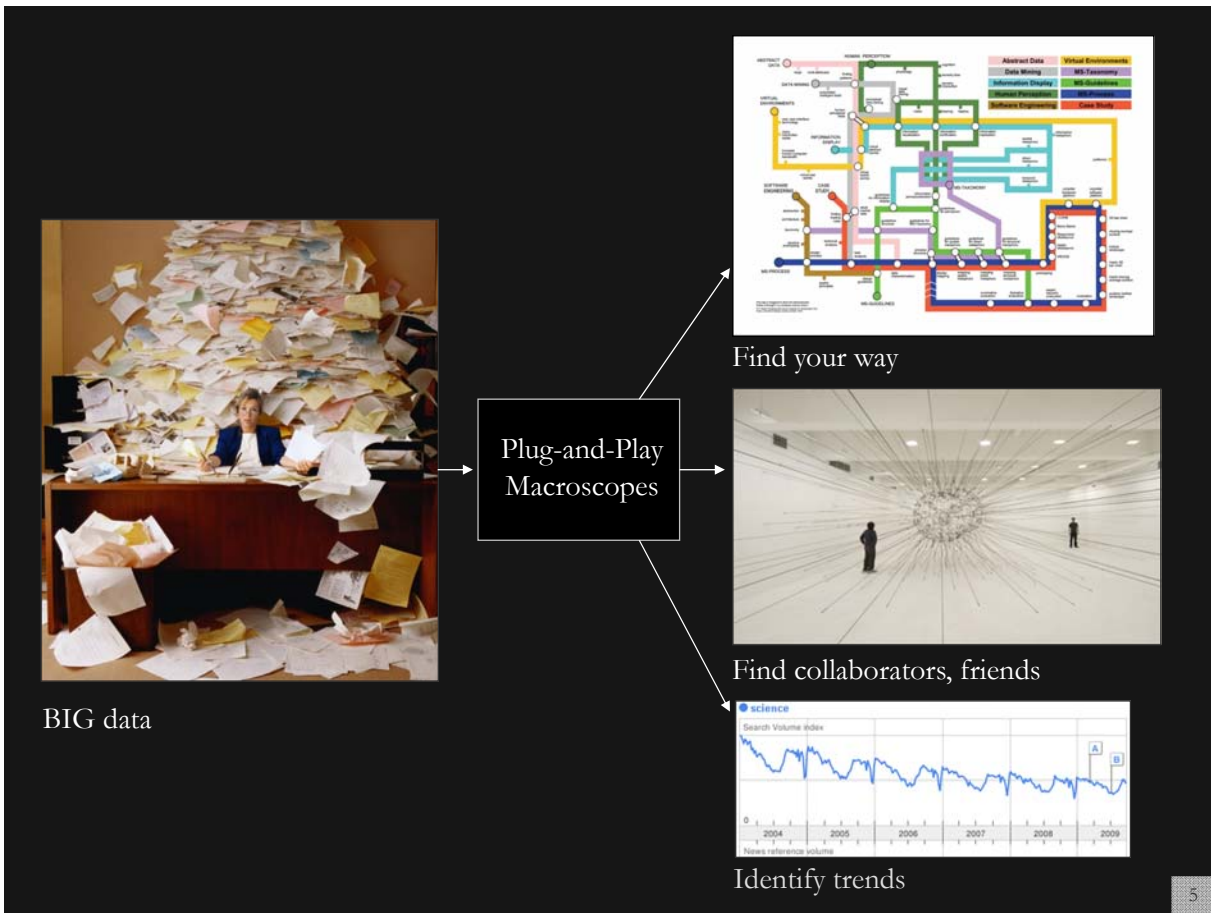
1. **Data mining and visualization research** that aims to increase our scientific understanding of the structure and dynamics of science and technology.
2. **Novel approaches and services** that improve information access, researcher networking, and research management.
3. **Data services and plug-and-play macroscope tools** that commoditize data mining and visualization.



3




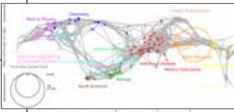
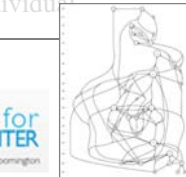
4



## Type of Analysis vs. Level of Analysis

	<i>Micro/Individual (1-100 records)</i>	<i>Meso/Local (101-10,000 records)</i>	<i>Macro/Global (10,000 &lt; records)</i>
<b>Statistical Analysis/Profiling</b>	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
<b>Temporal Analysis (When)</b>	Funding portfolio of one individual	Mapping topic bursts in 20-years of PNAS	113 Years of Physics Research
<b>Geospatial Analysis (Where)</b>	Career trajectory of one individual	Mapping a states intellectual landscape	PNAS publications
<b>Topical Analysis (What)</b>	Base knowledge from which one grant draws.	Knowledge flows in Chemistry research	VxOrd/Topic maps of NIH funding
<b>Network Analysis (With Whom?)</b>	NSF Co-PI network of one individual	Co-author network	NIH's core competency

## Type of Analysis vs. Level of Analysis

	<i>Micro/Individual</i> (1-100 records)	<i>Meso/Local</i> (101-10,000 records)	<i>Macro/Global</i> (10,000 < records)
<b>Statistical Analysis/Profiling</b>	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
<b>Temporal Analysis (When)</b>	Funding portfolio of one individual	Mapping topic bursts in 20-years of PNAS	113 Years of Physics Research
<b>Geospatial Analysis (Where)</b>	Career trajectory of one individual 	Mapping a states intellectual landscape	PNAS publications
<b>Topical Analysis (What)</b>	Base knowledge flows. 	Knowledge flows in Chemistry research	VxOrd/Topic maps of NIH funding
<b>Network Analysis (With Whom?)</b>	Network of one individual 	Co-author network	NIH's core competency



7

## MICRO: Mapping Richard M. Shiffrin's Career

by by Katy Börner, Tedd Polley and Michael J. Stamper

### Time and Space:

1968 Ph.D. in Mathematical Psychology, Stanford University

1968 joins Faculty of the Department of Psychology and Brain Sciences, IUB

1997 Honorary doctorate from the University of Amsterdam

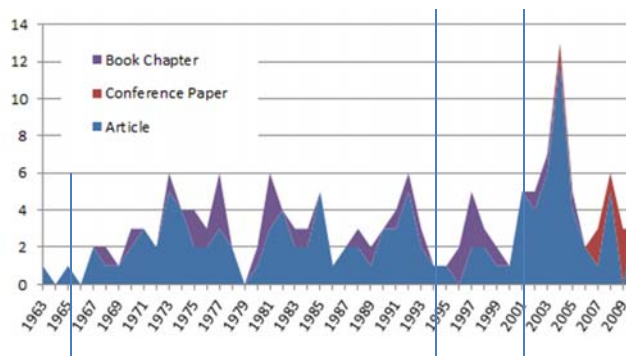
2004 Distinguished Scientific Contribution Award, American Psychological Association

### Topical Focus:

Modeling of human cognition in areas ranging from perception to attention to learning. Best known for explicit models of human memory.

### Dataset:

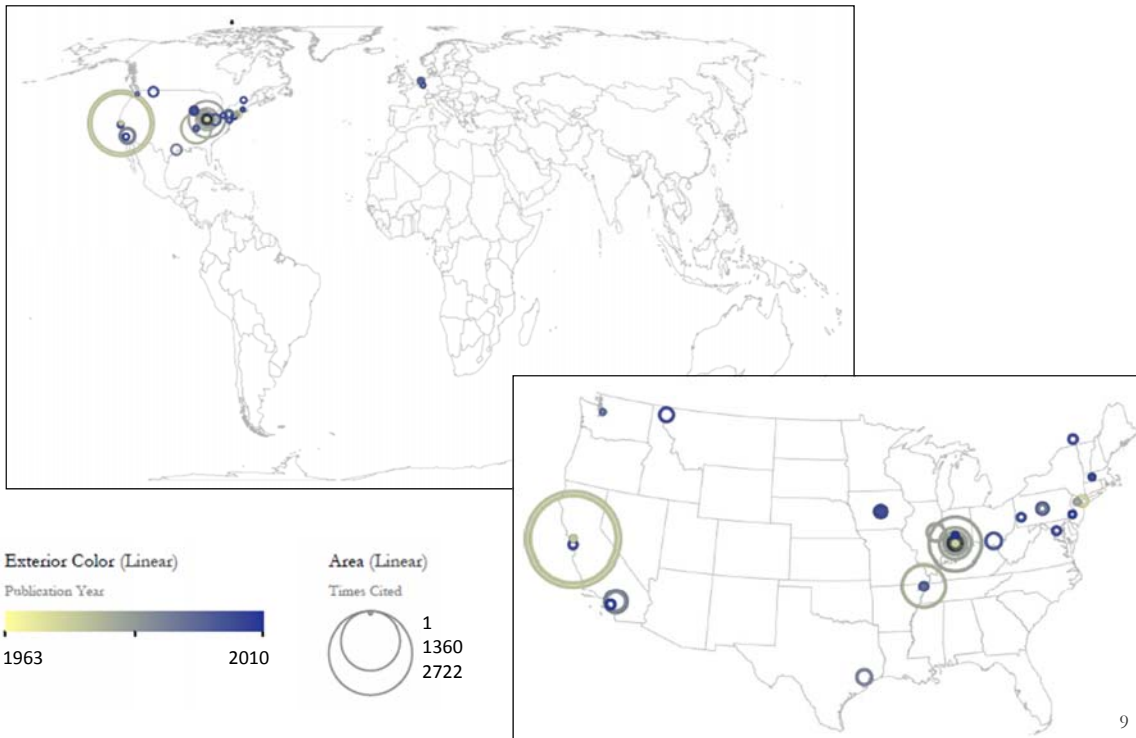
154 publications downloaded from ISI Web of Science plus Book chapters added by hand from resume (117 journal articles, 30 book chapters and 7 conference papers) for the years 1963-2010.



8

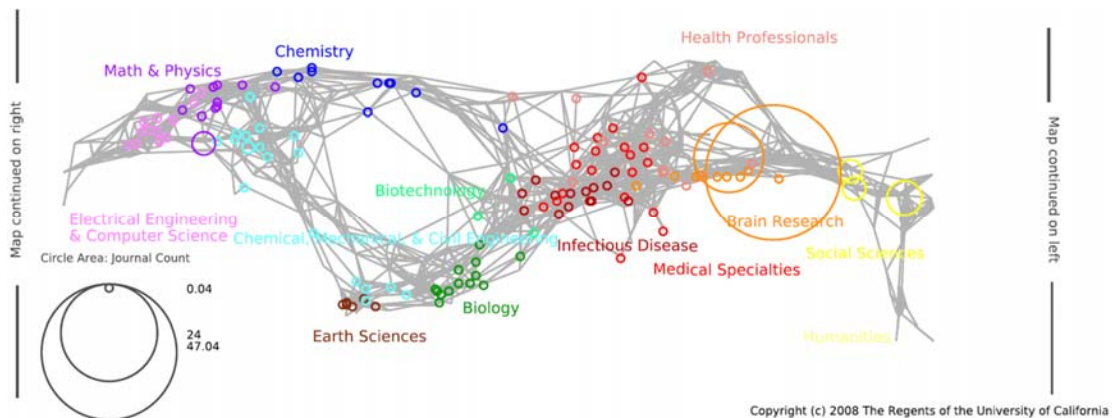


# GEO: Paper First Author Affiliations (World and zoom into U.S.)



# TOPICAL Distribution of Papers on UCSD Map of Science

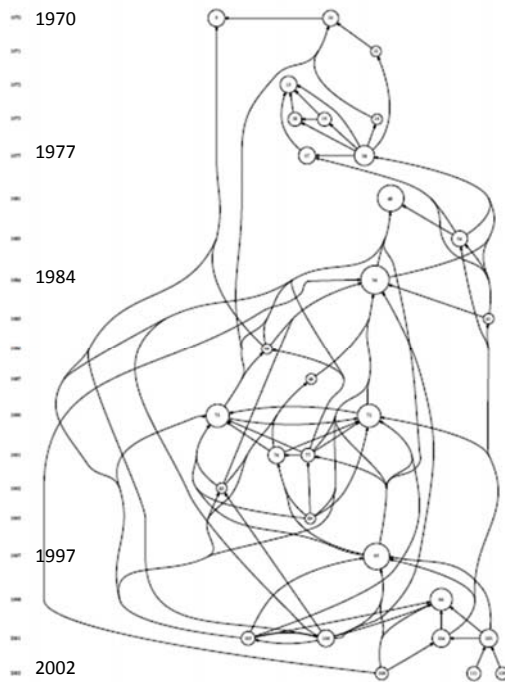
74 of 154 publications can be science located—fractionally assigned to 554 subdisciplines.



Copyright (c) 2008 The Regents of the University of California

- **Brain Research**
  - 1 American Journal Of Psychology
  - 1 Behavior Research Methods
  - 2 Behavioral And Brain Sciences
  - 1 Cognition
  - 3 Cognitive Psychology
  - 2 Cognitive Science
  - 3 Journal Of Experimental Psychology: General
  - 3 Journal Of Experimental Psychology: Human Perception And Performance
  - 19 Journal Of Experimental Psychology: Learning Memory And Cognition
  - 5 Journal Of Mathematical Psychology
  - 3 Journal Of Vision
  - 1 Perception
  - 1 Psychological Research
- **Math & Physics**
  - 1 Scientific American
- **Social Sciences**
  - 1 Annual Review Of Psychology
  - 3 International Journal Of Psychology
  - 1 Journal Of Personality And Social Psychology
- **Multiple Categories**
  - 3 Journal Of The Acoustical Society Of America
  - 2 Proceedings Of The National Academy Of Sciences Of The United States Of Ame...
  - 4 Science
- **Psychological Review**
  - 1 Quarterly Journal Of Experimental Psychology Section A: Human Experimental ...
  - 1 Spatial Vision
  - 1 Vision Research

## Network: Paper Intercitations (Historiograph)



Only 30 papers that are cited more than two times within the set of 154 papers are included.

Early and recent papers have a hard time to acquire 3+ citations within the set.

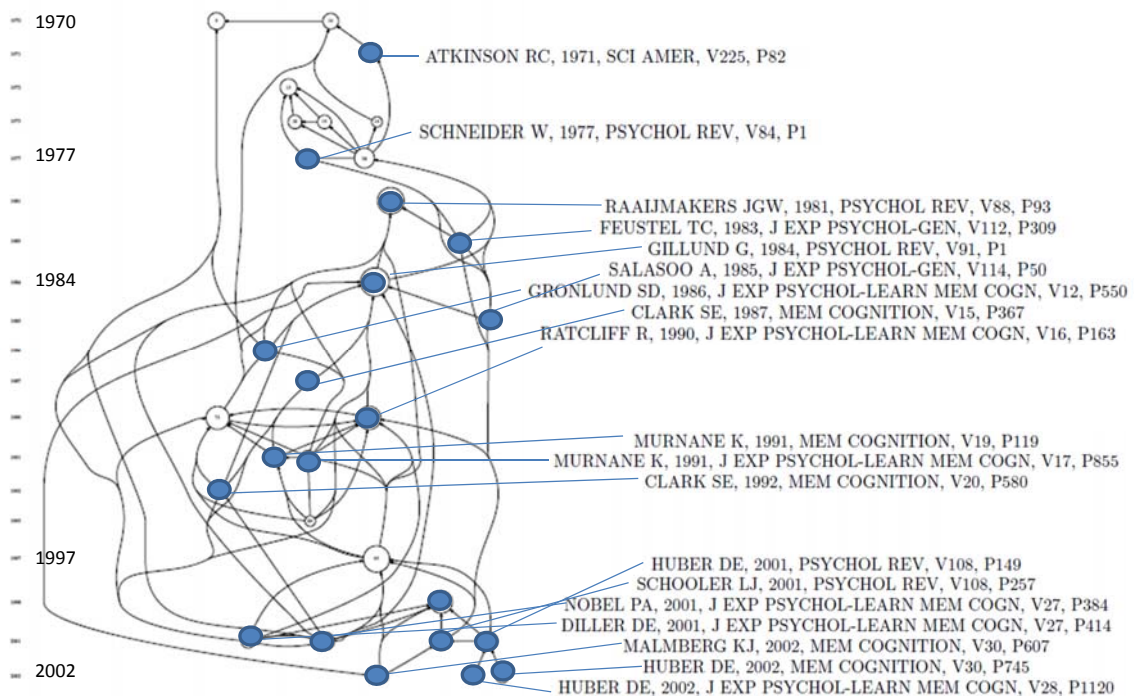
All papers that have a first author other than Shiffrin are labeled.

Circle size represents #citations.

Early papers highlight the “base knowledge” Shiffrin is drawing from. Later papers show Shiffrin’s impact on the next generation.

11

## Network: Paper Intercitations (Historiograph)



12

## Network: Paper Intercitations (Historiograph)

ID	Citation	LCS	GCS
9	SHIFFRIN RM, 1970, MODELS OF MEMORY, P375	7	
10	SHIFFRIN RM, 1970, SCIENCE, V168, P1601	6	45
13	ATKINSON RC, 1971, SCI AMER, V225, P82	3	336
15	SHIFFRIN RM, 1972, J EXP PSYCHOL, V93, P72	7	166
16	SHIFFRIN RM, 1973, J EXP PSYCHOL, V100, P39	3	46
18	SHIFFRIN RM, 1973, PERCEPT PSYCHOPHYS, V13, P328	5	35
19	SHIFFRIN RM, 1973, PERCEPT PSYCHOPHYS, V14, P231	5	24
37	SCHNEIDER W, 1977, PSYCHOL REV, V84, P1	7	2232
38	SHIFFRIN RM, 1977, PSYCHOL REV, V84, P127	10	2722
48	RAAIJMAKERS JGW, 1981, PSYCHOL REV, V88, P93	18	566
54	FEUSTEL TC, 1983, J EXP PSYCHOL-GEN, V112, P309	6	294
56	GILLUND G, 1984, PSYCHOL REV, V91, P1	20	834
61	SALASOO A, 1985, J EXP PSYCHOL-GEN, V114, P50	3	147
64	GRONLUND SD, 1986, J EXP PSYCHOL-LEARN MEM COGN, V12, P550	3	40
66	CLARK SE, 1987, MEM COGNITION, V15, P367	3	23
72	RATCLIFF R, 1990, J EXP PSYCHOL-LEARN MEM COGN, V16, P163	14	150
73	SHIFFRIN RM, 1990, J EXP PSYCHOL-LEARN MEM COGN, V16, P179	14	106
76	MURNANE K, 1991, MEM COGNITION, V19, P119	7	35
77	MURNANE K, 1991, J EXP PSYCHOL-LEARN MEM COGN, V17, P855	5	52
83	CLARK SE, 1992, MEM COGNITION, V20, P580	3	39
89	SHIFFRIN RM, 1995, J EXP PSYCHOL-LEARN MEM COGN, V21, P267	3	88
95	Shiffirin RM, 1997, PSYCHONOMIC BULL REV, V4, P145	17	284
96	SHIFFRIN RM, 1998, RATIONAL MODELS OF COGNITION, P73	12	
103	HUBER DE, 2001, PSYCHOL REV, V108, P149	8	39
104	SCHOOLER LJ, 2001, PSYCHOL REV, V108, P257	8	32
105	NOBEL PA, 2001, J EXP PSYCHOL-LEARN MEM COGN, V27, P384	5	39
106	DILLER DE, 2001, J EXP PSYCHOL-LEARN MEM COGN, V27, P414	7	29
108	MALMBERG KJ, 2002, MEM COGNITION, V30, P607	4	32
109	HUBER DE, 2002, MEM COGNITION, V30, P745	4	18
111	HUBER DE, 2002, J EXP PSYCHOL-LEARN MEM COGN, V28, P1120	5	17

**LCS**  
Local Citation  
Score – cites  
within  
dataset

**GCS**  
Global  
Citation  
Score – all  
cites in Web  
of Science

Records with  
missing GCS  
are not in  
WoS.

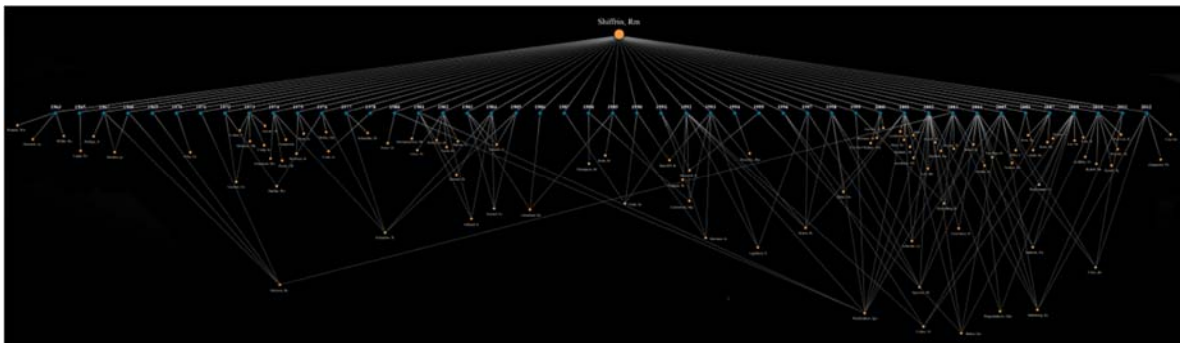
13

## Network: Evolving Collaborations

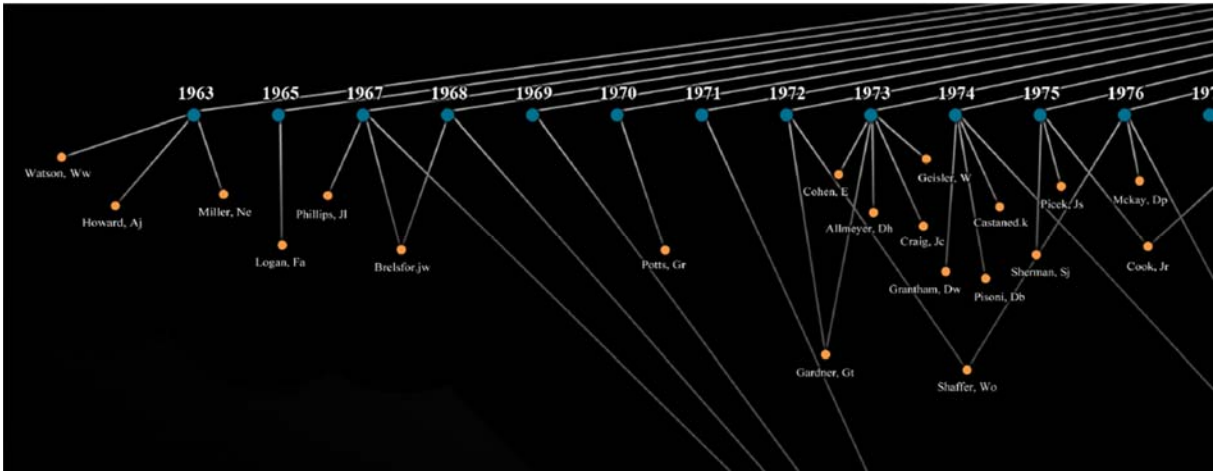
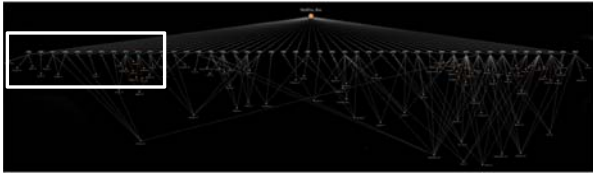
Shown here is a bimodal network of all 89 unique authors (in orange) connected to years in which they co-authored publications (in blue).

Basic author unification was performed to ensure there is exactly one author node for Shiffrin R M, Shiffrin, R, Shiffrin, Richard, Shiffrin, Richard M. Shiffrin, RM and each of his co-authors.

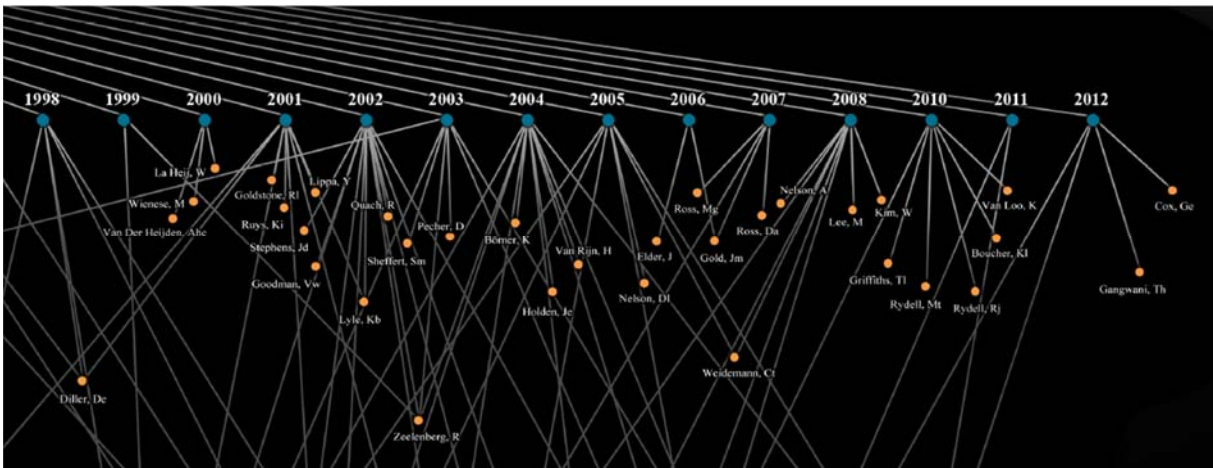
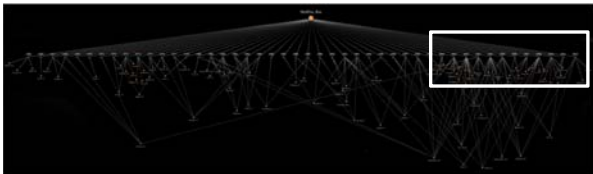
Large at [http://cns.iu.edu/research/2012-shiffrin\\_map-40x36.pdf](http://cns.iu.edu/research/2012-shiffrin_map-40x36.pdf)



14



15



16

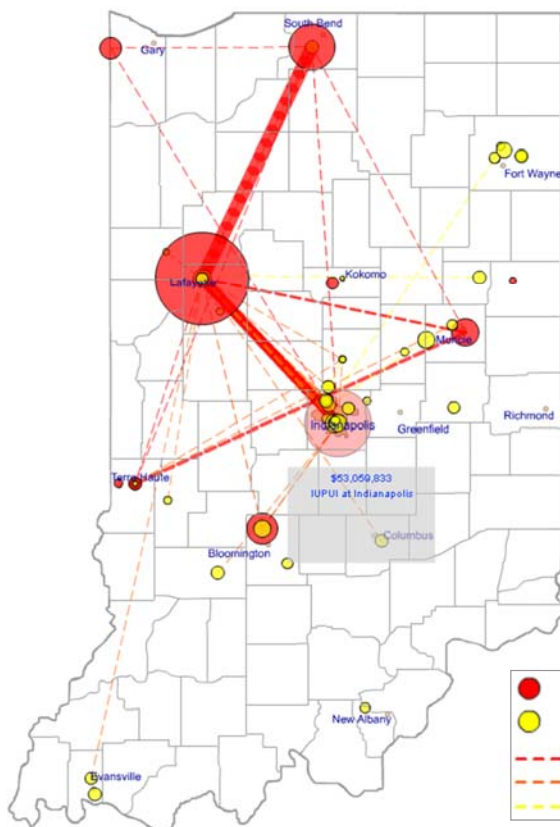


## Type of Analysis vs. Level of Analysis

	<i>Micro/Individual</i> (1-100 records)	<i>Meso/Local</i> (101-10,000 records)	<i>Macro/Global</i> (10,000 < records)
<b>Statistical Analysis/Profiling</b>	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
<b>Temporal Analysis (When)</b>	Funding portfolio of one individual	Mapping topic bursts in 20-years of PNAS	113 Years of Physics Research
<b>Geospatial Analysis (Where)</b>	Career trajectory of one individual	Mapping a state intellectual landscape	PNAS publications
<b>Topical Analysis (What)</b>	Base knowledge from which one grant draws.	Knowledge from Chemistry research	VxOrd/Topic maps of NIH funding
<b>Network Analysis (With Whom?)</b>	NSF Co-PI network of one individual	NSF Co-PI network	NIH's core competency



17



## MESO: Geo

### Mapping Indiana's Intellectual Space

Identify

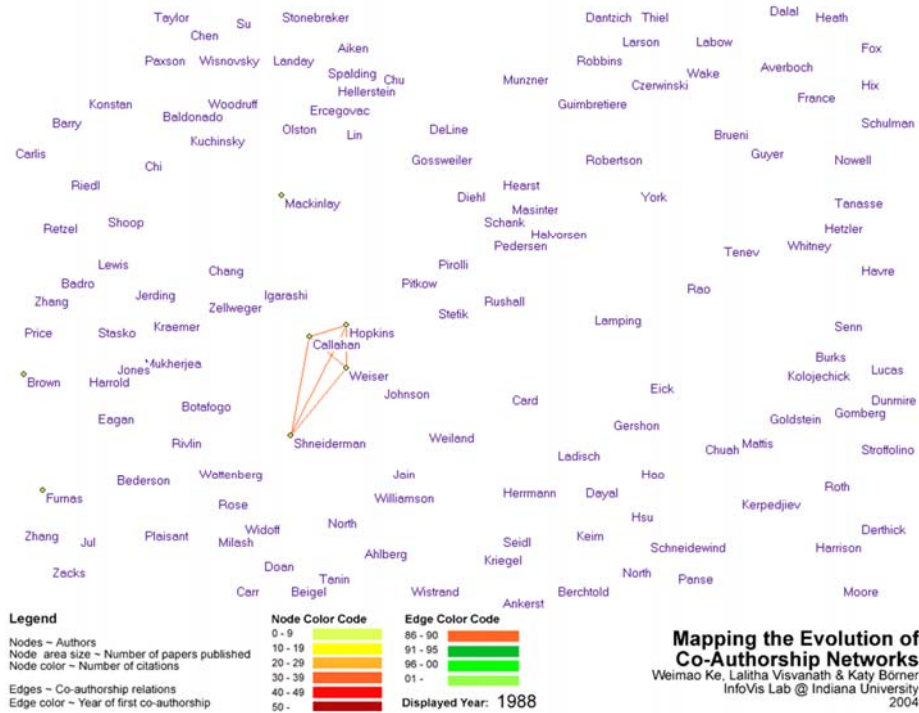
- Pockets of innovation
- Pathways from ideas to products
- Interplay of industry and academia

18

# MESO: Network

## Mapping the Evolution of Co-Authorship Networks

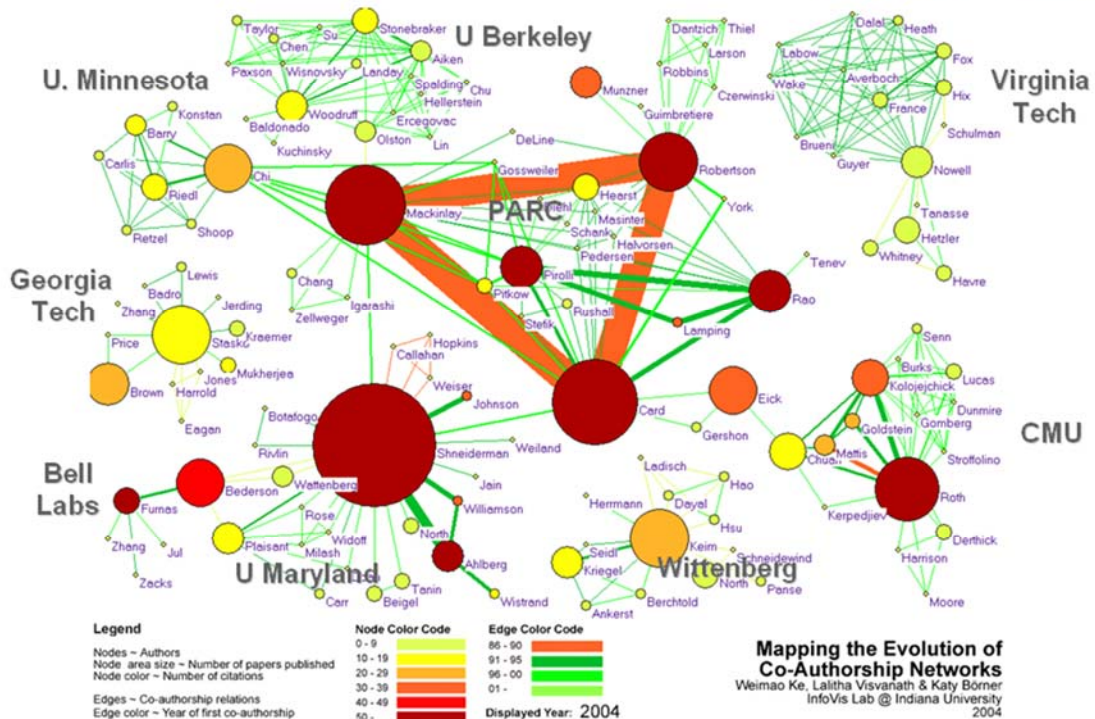
Ke, Visvanath & Börner, (2004) Won 1st price at the IEEE InfoVis Contest.



19

## Mapping the Evolution of Co-Authorship Networks

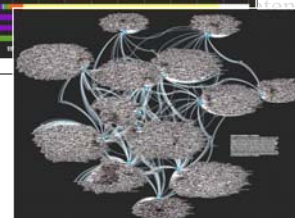
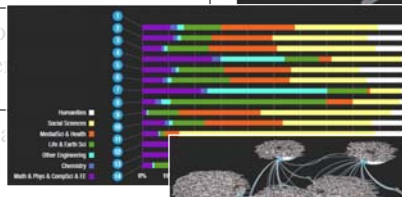
Ke, Visvanath & Börner, (2004) Won 1st price at the IEEE InfoVis Contest.



20

## Type of Analysis vs. Level of Analysis

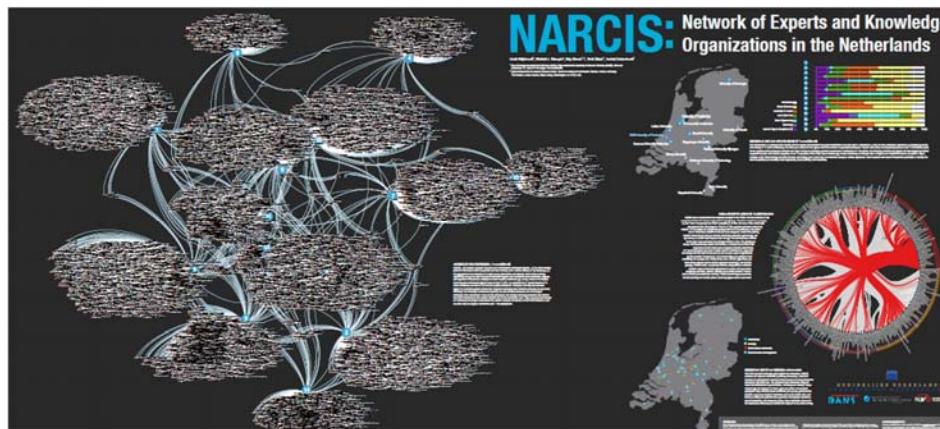
	<i>Micro/Individual</i> (1-100 records)	<i>Meso/Local</i> (101-10,000 records)	<i>Macro/Global</i> (10,000 < records)
<b>Statistical Analysis/Profiling</b>	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
<b>Temporal Analysis (When)</b>	Funding portfolio of one individual	Mapping topic bursts in 20-years of PNAS	
<b>Geospatial Analysis (Where)</b>	Career trajectory of one individual	Mapping a states intellectual landscape	
<b>Topical Analysis (What)</b>	Base knowledge from which one grant draws.	Knowledge in Chemistry	
<b>Network Analysis (With Whom?)</b>	NSF Co-PI network of one individual	Co-PI network	



## MACRO: Mapping NARCIS

by Linda Reijnhoudt, Michael J. Stamper, Katy Börner, Chris Baars, Andrea Scharnhorst

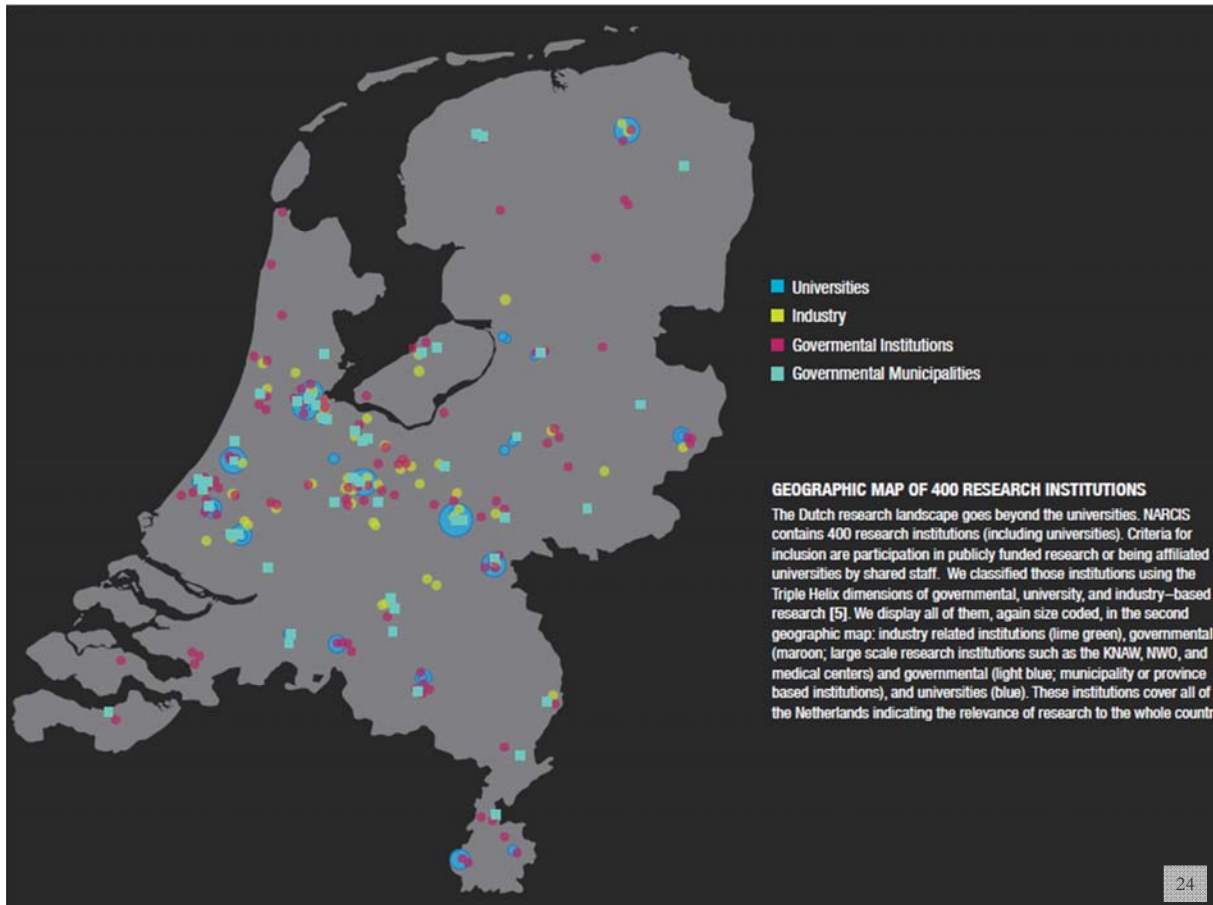
NARCIS—the National Academic Research and Collaborations Information System—is the Dutch national portal for information about researchers and their scholarly work. It consists of two databases: (1) An aggregation of the metadata of all open access publications and datasets within Dutch scholarly repositories and (2) the Dutch Research Information System with information about institutions, projects, and expertise of researchers employed at Dutch research institutions. All analyses use a dump of the NARCIS Dutch Research Information System made on April 3, 2012.



# Geo



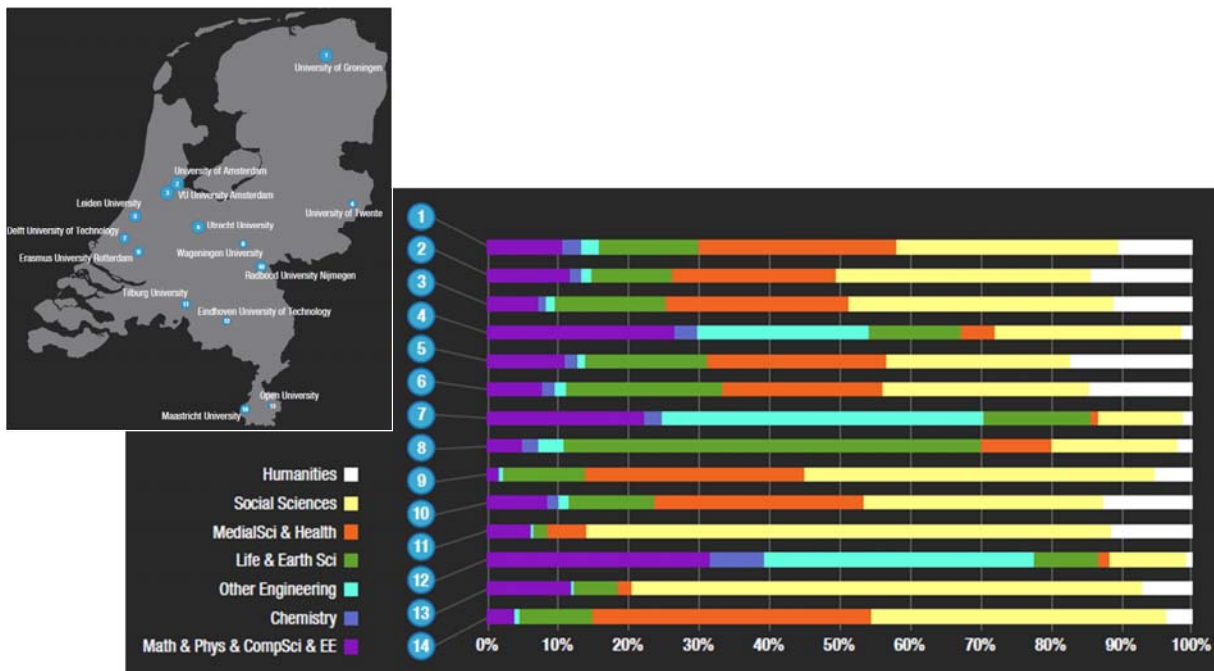
23



24

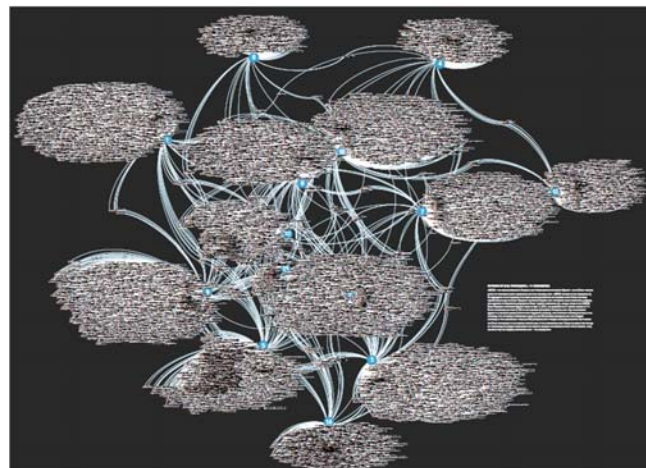


# Topical: Expertise Profiles

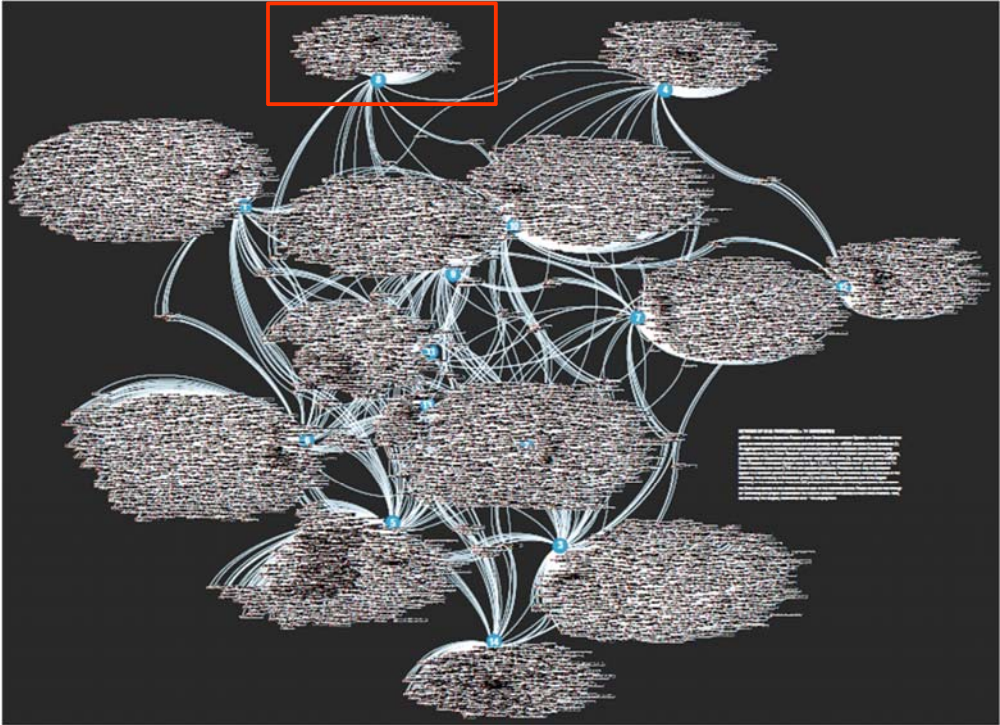


# Network: Faculty-University Affiliations

Shown is the bi-partite network of 14 Dutch universities and their 8,156 full and assistant professors. Professors with only one affiliation appear in a dandelion pattern around their respective university, their multitude indicating the size of the university. The University of Amsterdam (Node 2) has 932 affiliated professors and is the largest university in The Netherlands. There are 265 professors with multiple university affiliations that interlink the Dutch universities to one national network also on the institutional level.

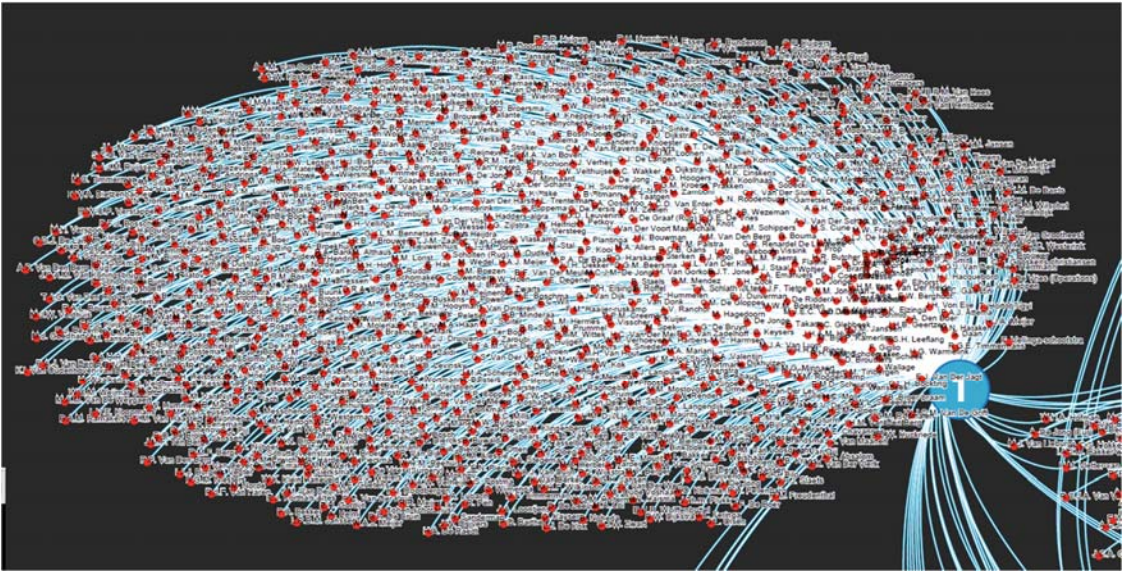


# Network: Faculty-University Affiliations



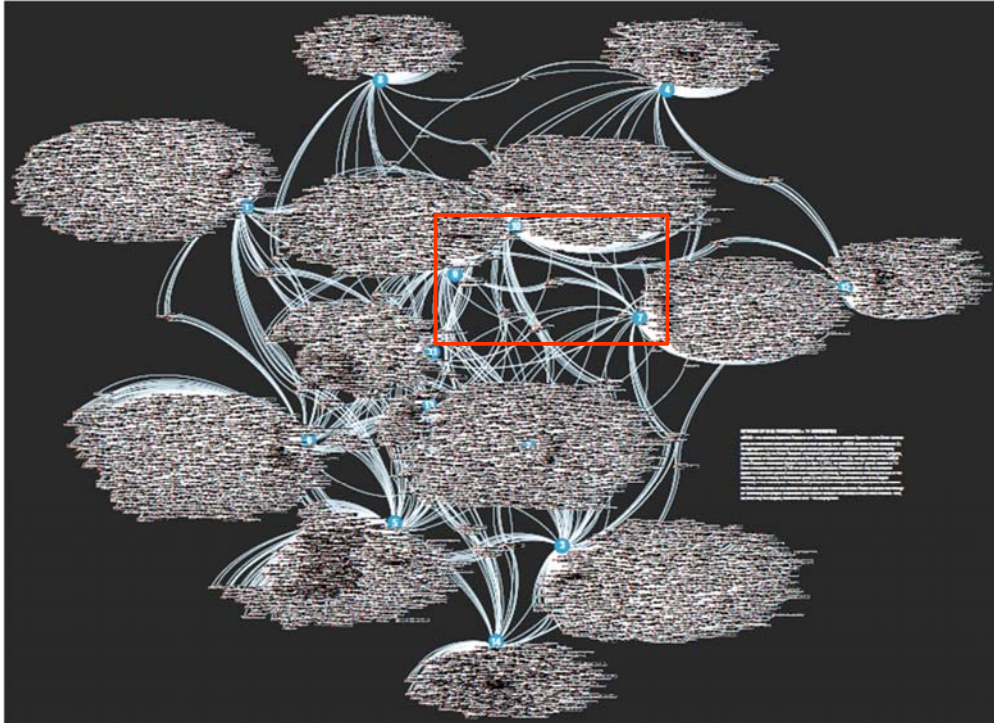
# Network: Faculty-University Affiliations

zoom into University of Groningen (1)





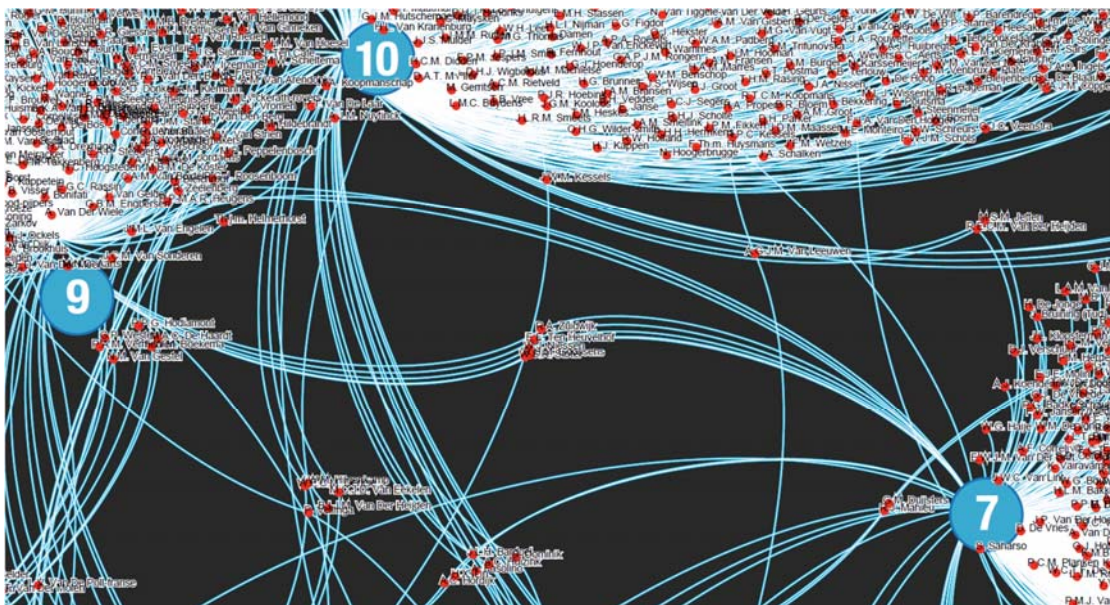
# Network: Faculty-University Affiliations



29

# Network: Faculty-University Affiliations

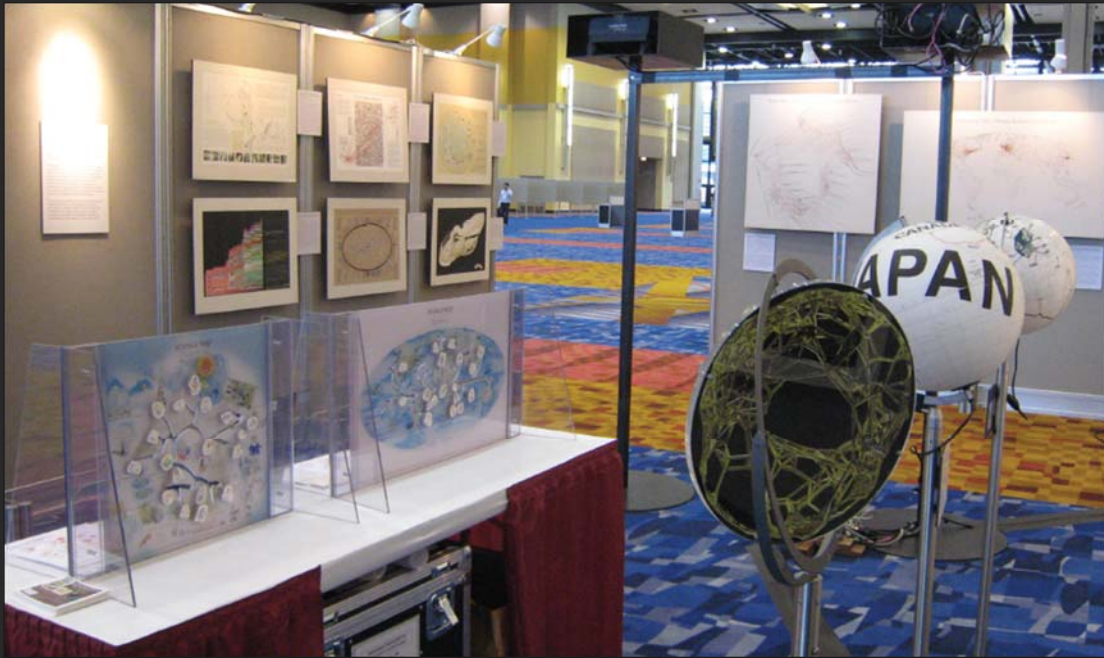
Zoom into Radboud University Nijmegen (10), Erasmus University Rotterdam (9), and Delft University of Technology (7) to see faculty with multiple affiliations.



30

## Mapping Science Exhibit – 10 Iterations in 10 years

<http://scimaps.org/>



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>



# The EMERGENCE of NANOTECHNOLOGY

## MAPPING THE NANO REVOLUTION

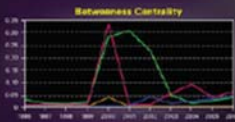
The emergence of nanotechnology has been one of the major scientific/technological revolutions in the last decade and it led to a structural reorganization of major fields of science. Poiné (1992) showed that fields of science and their development can be mapped using aggregated citations among the journals in the fields and their relevant interconnections.

The frames to the right show the evolving journal citation networks for the years 1998-2003. Distances are proportional to cosine values between the citation patterns of the respective journals. The usual destinations of key articles during the development of nanotechnology are given below each frame. Most notably, leading papers in Science and Nature catalyzed the breakthrough around 2000.

## CHANGING ROLES OF DIFFERENT JOURNALS

The interdisciplinarity of a journal can be measured using betweenness centrality (BC) – journals that occur on many shortest paths between other journals in a network have higher BC values than those that do not. In the maps, sizes of nodes are proportional to the betweenness centrality of the respective journal in the citation network.

From being a specialist journal in applied physics, the journal *Nanotechnology* obtains a high BC value in the years of its transition (ca. 2001). This is preceded by the "invention" of Science. After the transition, the new field of nanotechnology is established, new journals such as *Nano Letters* published by the influential American Chemical Society take the lead, and a new social structure with low BC value journals results.



An animated sequence of this evolution is at <http://www.levinscience.com>.

## References

Leydesdorff, L. and T. Schank, 2006. Dynamic Animations of Journal Maps: Indicators of Structural Change and Interdisciplinary Developments. *Journal of the American Society for Information Science and Technology*, 57(11), 1210-1218.

Poiné, Daniel J. de Souza (1992). Networks of scientific papers. *Science*, 149, no. 3662, 510-515.

**1998**

During the period 1998-2000, the journal *Nanotechnology* is published as a journal in *Applied Physics*.

**1999**

Technology-oriented journals play a role in the broader report arrangement of the journal *Nanotechnology*.

**2000**

The journal *Science* interfaces with research journals in both pure chemistry and applied physics. *Nanotechnology* emerges as core journal.

**2001**

The journal *Nanotechnology* now provides the interface between chemistry and physics. The "invention" by Science is no longer needed.

## LEGEND



Values  
0.0  
0.2  
0.33

**2003**

The journal *Science* is relevant in the citation impact structure, and new journals as one of the specialist journals in nanotechnology. *Nanotechnology* further develops as an increasingly integrated network of journals.

**2002**

Other journals in nanoscience and technology begin to emerge, and the leading role of the journal *Nanotechnology* gradually diminishes. *Nano Letters* and the *Journal of Nanoscience and Nanotechnology* join the new field of nanotechnology.

Loet Leydesdorff, Thomas Schank and the Journal of the American Society for Information Science and Technology, 2010. The Emergence of Nanoscience & Technology.

# DESIGN VS. EMERGENCE: VISUALIZATION OF KNOWLEDGE ORDERS

## WIKIPEDIA'S CATEGORY STRUCTURE

Wikipedia's category structure is a hierarchical tree of categories. Categories are used to group related articles. Categories are used to group related articles. Categories are used to group related articles. Categories are used to group related articles.

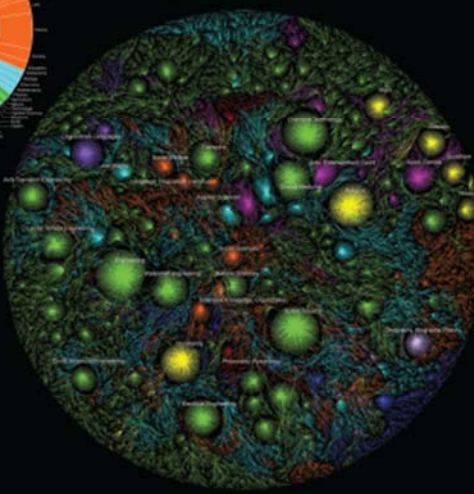


**CATEGORY DISTRIBUTION OF WIKIPEDIA & UDC**

This visualization shows the distribution of Wikipedia's categories across different levels of the hierarchy. The chart is a sunburst chart, where the inner ring represents the top-level categories, and the outer rings represent sub-categories. The colors represent different categories, and the size of the segments represents the number of articles in each category.

## UNIVERSAL DECIMAL CLASSIFICATION

The UDC (Universal Decimal Classification) is a hierarchical tree of categories. Categories are used to group related articles. Categories are used to group related articles. Categories are used to group related articles.



Almila Akdag Salah, Cheng Gao, Krzysztof Suchacki, and Andrea Scharnhorst (2011) Design vs. Emergence: Visualization of Knowledge Orders.



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>



Illuminated Diagram Display  
 soon on display at the Smithsonian  
 in DC.

[http://scimaps.org/exhibit\\_info/#ID](http://scimaps.org/exhibit_info/#ID)



### Geographic Map: Where Science Gets Done

### Science Map: How Scientific Disciplines Relate

**About**

This Illuminated Diagram display adds the flexibility of an interactive program to the incredibly high data density of a print. This technique is generally useful when there is too much pertinent data to be displayed on a screen but the data is relatively stable. The computer can direct the eye to what's important by using projectors or screens as smart spotlights, animating the research impact of individuals, giving a "grand tour" of science, or highlighting query results (as when you touch the lectern or use the keyboard) with an overlay of moving light.

**Elinor Ostrom - Nobel Prize in Economic Sciences 2009**

**Born:** 7 August 1933, New York, NY, USA

**Affiliation at the time of the award:** Indiana University, Bloomington, IN, USA, Arizona State University, Tempe, AZ, USA

**Prize motivation:** "for her analysis of economic governance, especially the commons"

**Field:** Economic governance

**Contribution:** Challenged the conventional wisdom by demonstrating how local property can be successfully managed by local commons without any regulation by central authorities or privatization.

**Interact**

Select any location on the Geographic Map location (by brushing your finger over an area on the lectern's touch screen) and topics studied in that area will highlight on the Science Map: the brighter a topic glows, the more papers on that topic originated in the selected area. Conversely, touching a scientific area in the Science Map illuminates places on the Geographic Map where that topic is studied. People and topic buttons support the exploration of publication output by selected Noble laureates and particular lines of research using MEDLINE data from 2000-2009.

Cancer	Cloning	HIV	Robert G. Edwards	Roger D. Kornberg	Elinor Ostrom
Obesity	Quality of Life	Smoking	Stanley B. Prusiner	Ahmed H. Zewail	View All

<http://scimaps.org>

**Keyword Search** 37

### Geographic Map: Where Science Gets Done

### Science Map: How Scientific Disciplines Relate

**About**

This Illuminated Diagram display adds the flexibility of an interactive program to the incredibly high data density of a print. This technique is generally useful when there is too much pertinent data to be displayed on a screen but the data is relatively stable. The computer can direct the eye to what's important by using projectors or screens as smart spotlights, animating the research impact of individuals, giving a "grand tour" of science, or highlighting query results (as when you touch the lectern or use the keyboard) with an overlay of moving light.

**Top Five Continents**

- North America - 4,000 records
- South & East Asia - 3,589
- Australia - 2,431
- Africa - 2,208
- South America - 1,562

**Top Five Scientific Disciplines**

- Math & Physics - 4,000 records
- Health Professions - 3,589
- Social Sciences - 2,431
- Aeronautical, Chemical, Mechanical & Civil Engineering - 2,208
- Humanities - 1,562

Input your search query here.

Q	W	E	R	T	Y	U	I	O	P
A	S	D	F	G	H	J	K	L	"
Z	X	C	V	B	N	M			
Space									Go

**Search**

The keyboard supports retrieval and display of papers based on their Medical Subject Headings (MeSH) and MeSH qualifier terms. If multiple terms are entered in a field, they are automatically combined using "OR". So, "breast cancer" matches any record with "breast" or "cancer" in that field. You can put AND between terms to combine with "AND". Thus "breast AND cancer" would only match records that contain both terms. Double quotation can be used to match compound terms, e.g., "breast cancer" retrieves records with the phrase "breast cancer", and not records where "breast" and "cancer" are both present, but the exact phrase.

<http://scimaps.org>

**People & Topics** 38



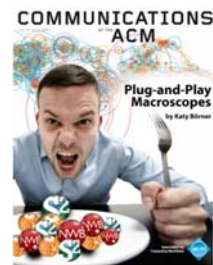
Ingo Gunther's Worldprocessor globe design now on display at the Giant Geo Cosmos OLED Display at the Museum of Emerging Science and Innovation in Tokyo, Japan





## Overview

1. **Data mining and visualization research** that aims to increase our scientific understanding of the structure and dynamics of science and technology.
2. **Novel approaches and services** that improve information access, researcher networking, and research management.
3. **Data services and plug-and-play macroscope tools** that commoditize data mining and visualization.



41

## Different Stakeholder Groups and Their Needs

### 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.

42

## Scholars Have Different Roles/Needs

**Researchers and Authors**—need to select promising research topics, students, collaborators, and publication venues to increase their reputation. They benefit from a global view of competencies, reputation and connectivity of scholars; hot and cold research topics and bursts of activity, and funding available per research area.

**Editors**—have to determine editorial board members, assign papers to reviewers, and ultimately accept or reject papers. Editors need to know the position of their journals in the evolving world of science. They need to advertise their journals appropriately and attract high-quality submissions, which will in turn increase the journal's reputation.

**Reviewers**—read, critique, and suggest changes to help improve the quality of papers and funding proposals. They need to identify related works that should be cited or complementary skills that authors might consider when selecting project collaborators.

**Teachers/Mentors**—teach classes, train doctoral students, and supervise postdoctoral researchers. They need to identify key works, experts, and examples relevant to a topic area and teach them in the context of global science.

**Inventors**—create intellectual property and obtain patents, thus needing to navigate and make sense of research spaces as well as intellectual property spaces.

**Investigators**—scholars need funding to support students, hire staff, purchase equipment, or attend conferences. Here, research interests and proposals have to be matched with existing federal and commercial funding opportunities, possible industry collaborators and sponsors.

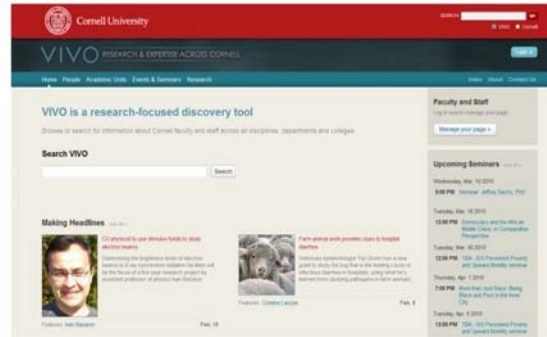
**Team Leads and Science Administrators**—many scholars direct multiple research projects simultaneously. Some have full-time staff, research scientists, and technicians in their laboratories and centers. Leaders need to evaluate performance and provide references for current or previous members; report the progress of different projects to funding agencies.



VIVO International Researcher  
Network

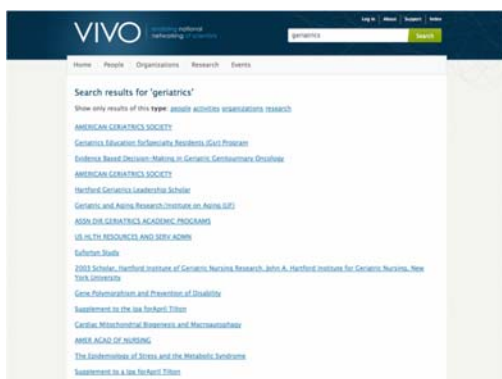
## VIVO: A Semantic Approach to Creating a National Network of Researchers (<http://vivoweb.org>)

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



Funded by \$12 million NIH award.

**Cornell University:** Dean Krafft (Cornell PI), Manolo Bevia, Jim Blake, Nick Cappadona, Brian Caruso, Jon Corson-Rikert, Elly Cramer, Medha Devare, John Ferreira, 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 Raam, 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 Ping, 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, Michaelen 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.





**University of Florida**

How do you want to compare?  
by Grants

Who do you want to compare?  
Search: X

Records 1 - 10 of 30

Entity Label	Grant Count	Entity Type
<input checked="" type="checkbox"/> Continuing Education	562	UF Department, Agent, Non-Academic Department, Department
<input checked="" type="checkbox"/> Florida Museum of Natural History	203	Museum, Agent
<input checked="" type="checkbox"/> College of Agricultural and Life Sciences	166	Agent, UF College, College
<input checked="" type="checkbox"/> College of Engineering	103	Agent, UF College, College
<input checked="" type="checkbox"/> Evelyn F. and William L. McKnight Brain Institute of the University of Florida	64	UF Center, Agent, Center
<input checked="" type="checkbox"/> International Center	54	UF Department, Agent, Non-Academic Department, Department
<input checked="" type="checkbox"/> Florida Sea Grant	44	UF Center, Agent, Center
<input type="checkbox"/> Whitney Laboratory for Marine Bioscience	42	UF Research Laboratory, Agent, Laboratory, Research Laboratory
<input type="checkbox"/> Water Institute	38	UF Center, Agent, Center
<input type="checkbox"/> College of Dentistry	35	Agent, UF College, College

Save as CSV Clear

**Comparing Grants of Organizations in University of Florida**

**Total Number of Grants**  
You have selected 7 of a maximum 10 organizations to compare. Clear

- Florida Sea Grant 44
- International Center 54
- Evelyn F. and William L. McKnight Brain Institute of the University of Florida 64
- College of Engineering 103
- College of Agricultural and Life Sciences 166
- Florida Museum of Natural History 203
- Continuing Education 562

**Temporal Analysis (When)** Temporal visualizations of the number of papers/funding award at the institution, school, department, and people level

**VIVO** enabling national networking of scientists

Home | People | Organizations | Research | Events

**University of Florida**

Explore 487 publications activity across 554 scientific sub-disciplines

13 Disciplines | 554 Sub-Disciplines

Search: X

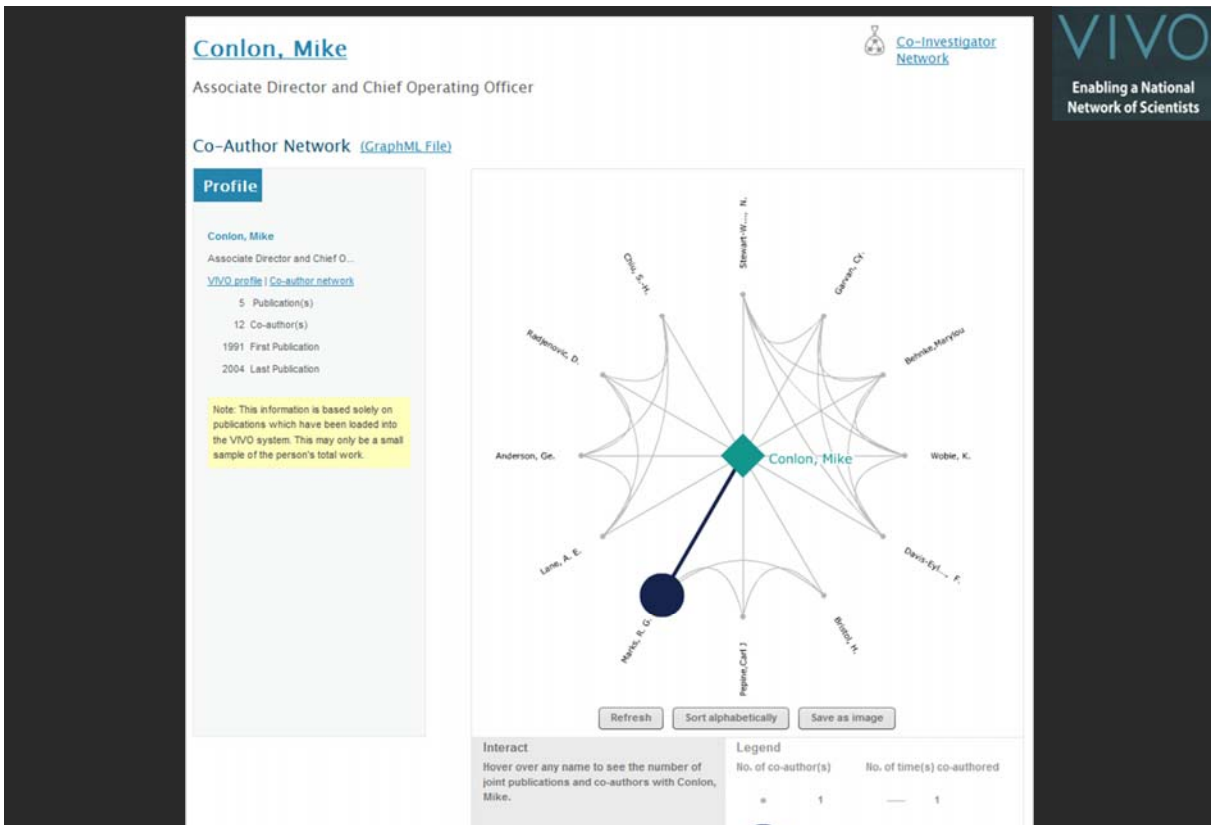
1 - 13 of 554

Sub-Disciplines	# of pubs.	% activity
Pest Management Science	24.2	5.0
Wildlife Research	19.1	3.9
Protein Science	13.1	2.7
Clinical Cancer Research	12.6	2.6
Pain	12.0	2.5
Environmental Contamination	11.2	2.3
Insect Physiology	11.1	2.3
Organic Chemistry	10.9	2.2
Marine Biology	10.3	2.1
Computer Aided Molecular Design	10.2	2.1
BioStatistics	9.0	1.9

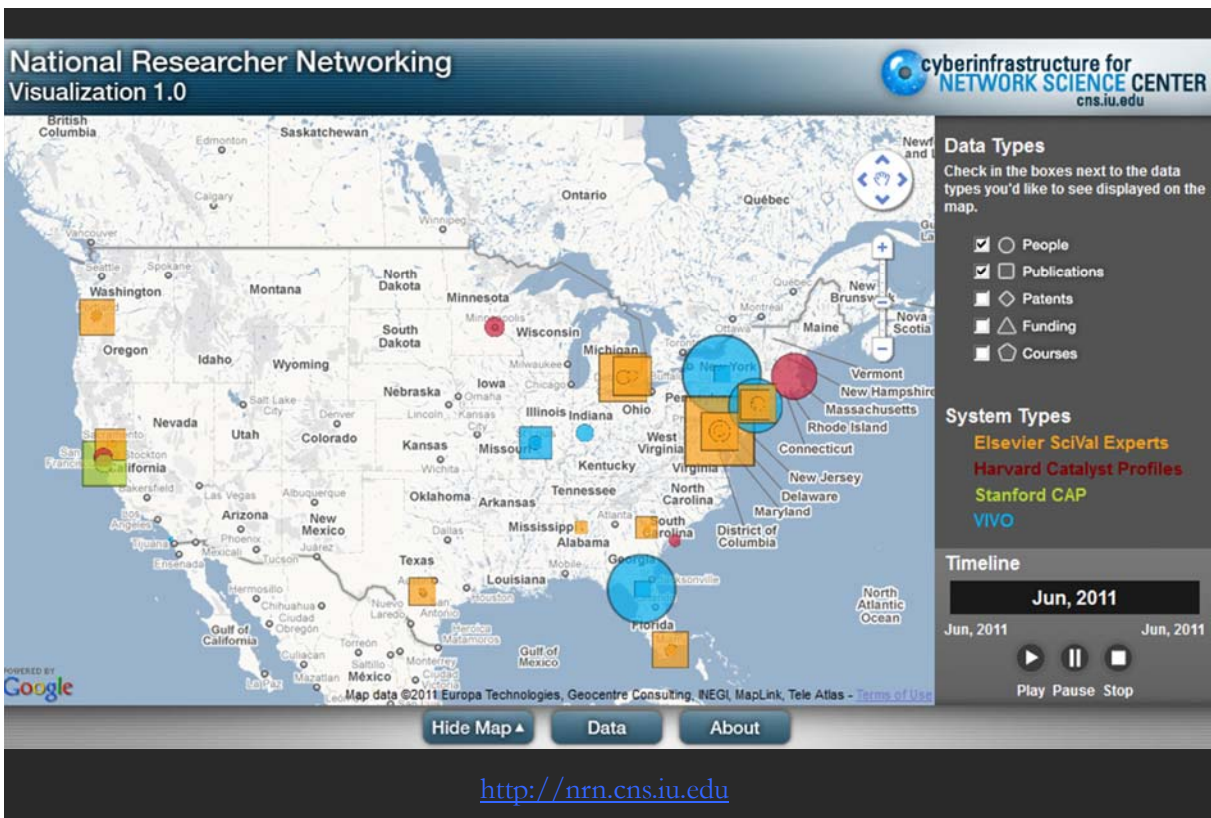
Top 290 disciplines shown

mapped 14.55% of 3,346 publications

**Topical Analysis (What)** Science map overlays will show where a person, department, or university publishes most in the world of science. (in work)



**Network Analysis (With Whom?)** Who is co-authoring, co-investigating, co-inventing with whom? What teams are most productive in what projects? 49



**Geospatial Analysis (Where)** Where is what science performed by whom? Science is global and needs to be studied globally. 50

## VIVO On-The-Go

## Overview, Interactivity, Details on Demand

come to  
commonly  
used devices  
and environments



## Develop VIVO Visualizations

See also *Visualization in VIVO Workshop on Aug 24, 2011*

<http://wiki.cns.in.edu/display/PRES/VIVO+Presentation>



### VIVO Presentation

4 Added by Chin Hua Kong, last edited by Chintan Tank on Aug 24, 2011 (view change)

### August, 2011 Workshop

#### Material

- [Java 1.5 or higher](#) - A programming language and computing platform for developing cross OS softwares.
- [Science of Science tool \(Sci2\)](#) - An desktop application for information analysis and visualization.
- [Gephi](#) - An interactive visualization tool for networks and complex systems, dynamic and hierarchical graphs.
- [VIVO August 2011 workshop data.zip](#) - Hands on workshop data package

#### Slides

- [Tutorial Slides](#) presented at the VIVO Conference 2011
- [Pre-Questionnaire](#) and [Post-Questionnaire](#)

#### Demo Links

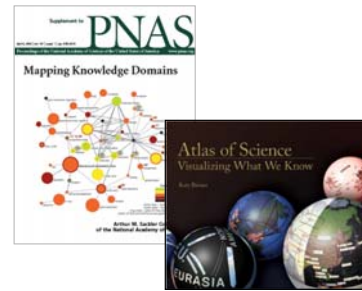
- [Map of Science Visualization \(dev link\)](#)
- [Temporal Graph Visualization \(dev link\)](#)
- [National Researcher Networking Visualization](#)
- [Word Cloud Visualization dev link](#)



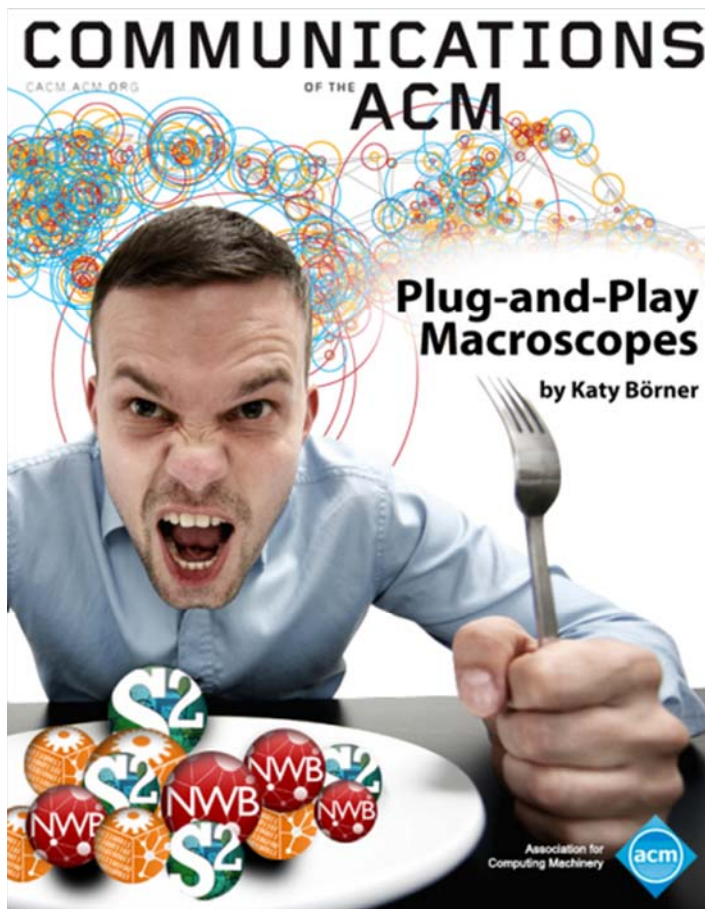


## Overview

1. **Data mining and visualization research** that aims to increase our scientific understanding of the structure and dynamics of science and technology.
2. **Novel approaches and data services** that improve information access, researcher networking, and research management.
3. **Data services and plug-and-play macroscope tools** that commoditize data mining and visualization.



55

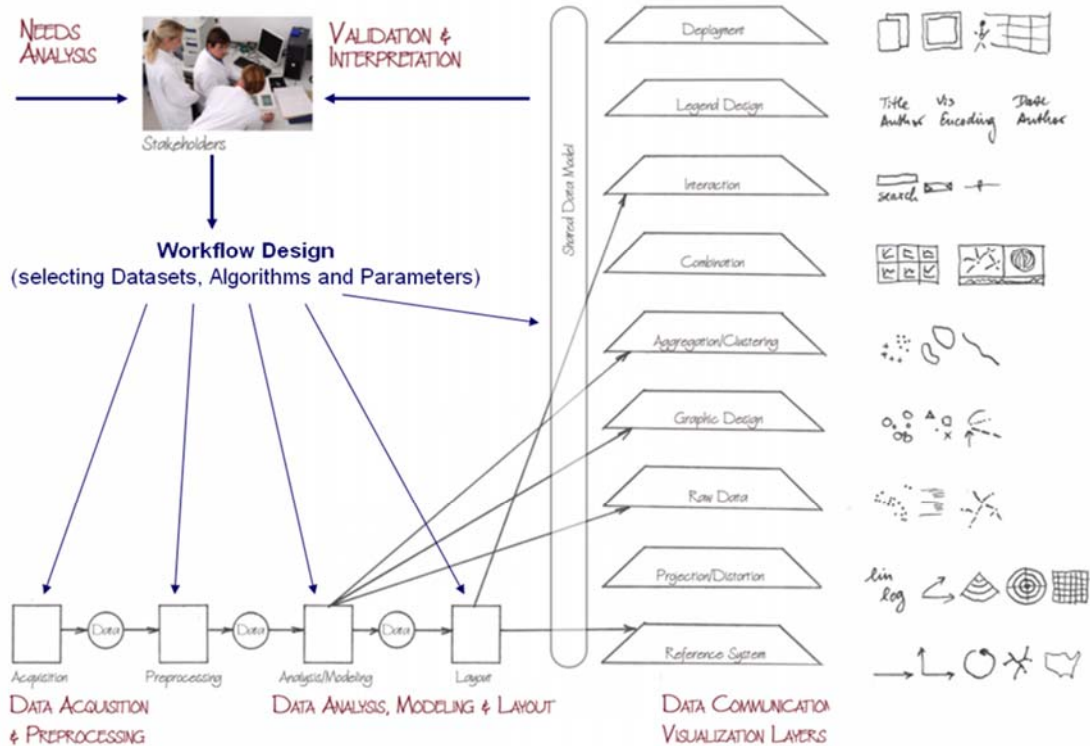


Börner, Katy. (March 2011). Plug-and-Play Macroscopes. *Communications of the ACM*, 54(3), 60-69.

Video and paper are at <http://www.scivee.tv/node/27704>

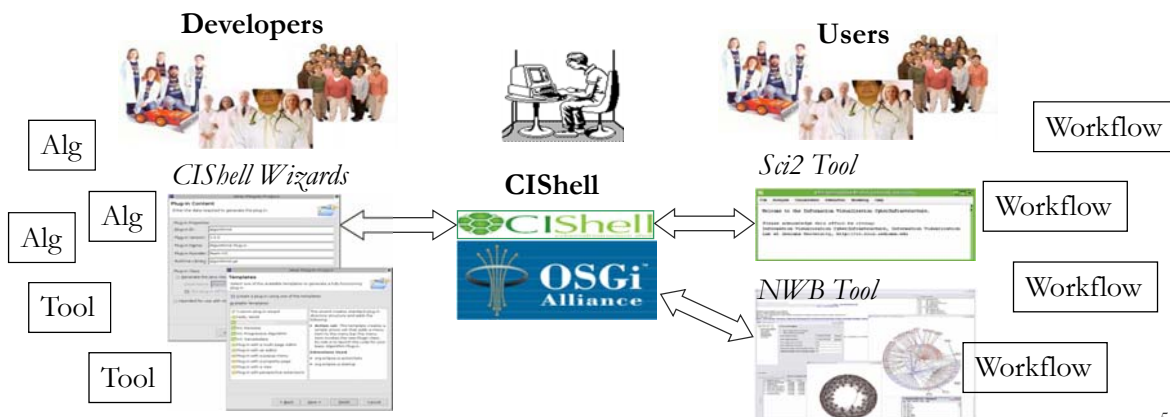
56

**Needs-Driven Workflow Design** using a modular data acquisition/analysis/ modeling/ visualization pipeline as well as modular visualization layers.



OSGi & CIShell

- CIShell (<http://cishell.org>) 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://osgi.org>), a standardized, component oriented, computing environment for networked services widely used in industry since more than 10 years.
- Specifically, CIShell provides “sockets” into which existing and new datasets, algorithms, and tools can be plugged using a wizard-driven process.







## CIShell Developer Guide

(<http://cishell.wiki.cns.iu.edu>)



Edit Add

1 Added by Micah Linnemeier, last edited by Micah Linnemeier on Mar 16, 2011 (view change)

### About the Cyberinfrastructure Shell

The Cyberinfrastructure Shell (CIShell) is an open source, community-driven platform for the integration and utilization of datasets, algorithms, tools, and computing resources. Algorithm integration support is built in for Java and most other programming languages. Being Java based, it will run on almost all platforms. The software and specification is released under an Apache 2.0 License.

CIShell is the basis of [Network Workbench](#), [TexTrend](#), [Sci²](#) and the upcoming [EpiC](#) tool.

CIShell supports remote execution of algorithms. A standard web service definition is in development that will allow pools of algorithms to transparently be used in a peer-to-peer, client-server, or web front-end fashion.

### CIShell Features

**A framework for easy integration of new and existing algorithms written in any programming language**

Using CIShell, an algorithm writer can fully concentrate on creating their own algorithm in whatever language they are comfortable with. Simple tools are provided to then take their algorithm and

### Learn More...

- [CIShell Papers](#)
- [CIShell Powered Tools](#)
- [Algorithms](#)
- [Plugins \(coming soon\)](#)
- [Misc. Tool Documentation](#)
- CIShell Web Services (coming soon)
- [Screenshots](#)

### Getting Started...

- [Documentation & Developer Resources](#)
- [Download](#)

### Getting Involved...

- [Contact Us](#)



## CIShell Portal (<http://cishell.org>)

**Cyberinfrastructure Shell (CIShell)**  
CIShell supports the plug-and-play of datasets and algorithms and their bundling into custom tools that serve the specific needs of a user group or research community. It has been applied to develop diverse custom tools, see below. Feel free to take plugins from any of these tools to design your personal dream tool.

Visit the **CIShell wiki** to learn more about using CIShell as a platform for your tool!

Provided by the [Cyberinfrastructure for Network Science Center](#) at Indiana University.

**Learn more about existing CIShell-powered tools below.**

**Network Workbench Tool (NWB)**  
The NWB Tool supports researchers, educators, and practitioners interested in the study of biomedical, social and behavioral science, physics, and other networks. It comes with a 77-page [user manual](#).

**Science of Science Tool (Sci²)**  
The Sci² Tool was specifically developed for science policy makers and researchers that study science by scientific means. It supports the temporal, geospatial, topical, and network analysis and visualization of scholarly datasets at the micro (individual), meso (local), and macro (global) levels. There exists a [112-page user manual](#) and 24 hours of [NIM tutorials](#) in this tool.

The Network Workbench (NWB) tool supports researchers, educators, and practitioners interested in the study of biomedical, social and behavioral science, physics, and other networks.

In February 2009, the tool provides more than 169 plugins that support the preprocessing, analysis, modeling, and visualization of networks.

**More than 50 of these plugins can be applied or were specifically designed for S&T studies.**

It has been downloaded more than 110,000 times since December 2006.



**Network Workbench**  
A Workbench for Network Science

Home People Research Publications

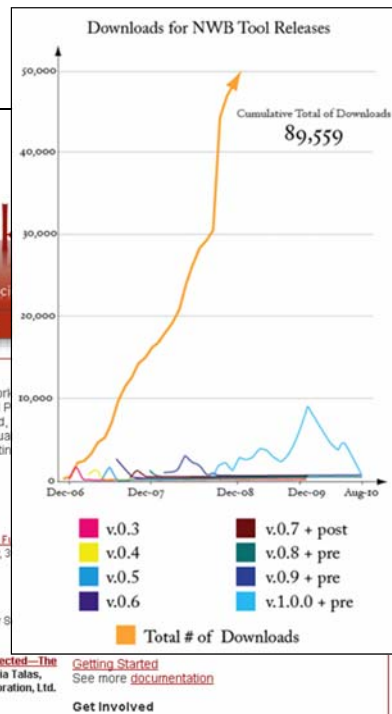
**Summary**  
Network Workbench: A Large-Scale Network Toolkit for Biomedical, Social Science and Psychology. Evaluate, and operate a unique distributed, scale network analysis, modeling, and visualization (NWB). The envisioned data-code-computer more  
[How to cite this project](#)

**News & Updates**

- 5.1.09 Kaelble, Steve. 2009. [Mapping the Frontiers of Knowledge, Research & Creative Activity, 3](#) (website accessed 5/1/09)
- 3.23.09 1.0.0 beta 5 Released
- 1.23.09 Ann Mcranie's [tutorial abstract](#) for S 2009
- 11.4.08 Two NWB PIs featured in "Connected—The Power of Six Degrees." 2008. Anna Maria Talas, Director. Australian Broadcasting Corporation, Ltd. [YouTube!](#) [Full Video](#) (300MB)

[Getting Started](#)  
See more [documentation](#)

[Get Involved](#)



Herr II, Bruce W., Huang, Weixia (Bonnie), Penumarthy, Shashikant & Börner, Katy. (2007). Designing Highly Flexible and Usable Cyberinfrastructures for Convergence. In Bainbridge, William S. & Roco, Mihail C. (Eds.), *Progress in Convergence - Technologies for Human Wellbeing* (Vol. 1093, pp. 161-179), *Annals of the New York Academy of Sciences*, Boston, MA.

61

## Computational Proteomics

What relationships exist between protein targets of all drugs and all disease-gene products in the human protein-protein interaction network?

Yildirim, Muhammed A., Kwan-II Goh, Michael E. Cusick, Albert-László Barabási, and Marc Vidal. (2007). *Drug-target Network*. *Nature Biotechnology* 25 no. 10: 1119-1126.



**Figure 2** Drug-target network (DT network). The DT network is generated by using the known associations between FDA-approved drugs and their target proteins. Circles and rectangles correspond to drugs and target proteins, respectively. A link is placed between a drug node and a target node if the protein is a known target of that drug. The area of the drug (protein) node is proportional to the number of targets that the drug has (the number of drugs targeting the protein). Color codes are given in the legend. Drug nodes (circles) are colored according to their Anatomical Therapeutic Chemical Classification, and the target proteins (rectangular boxes) are colored according to their cellular component obtained from the Gene Ontology database.

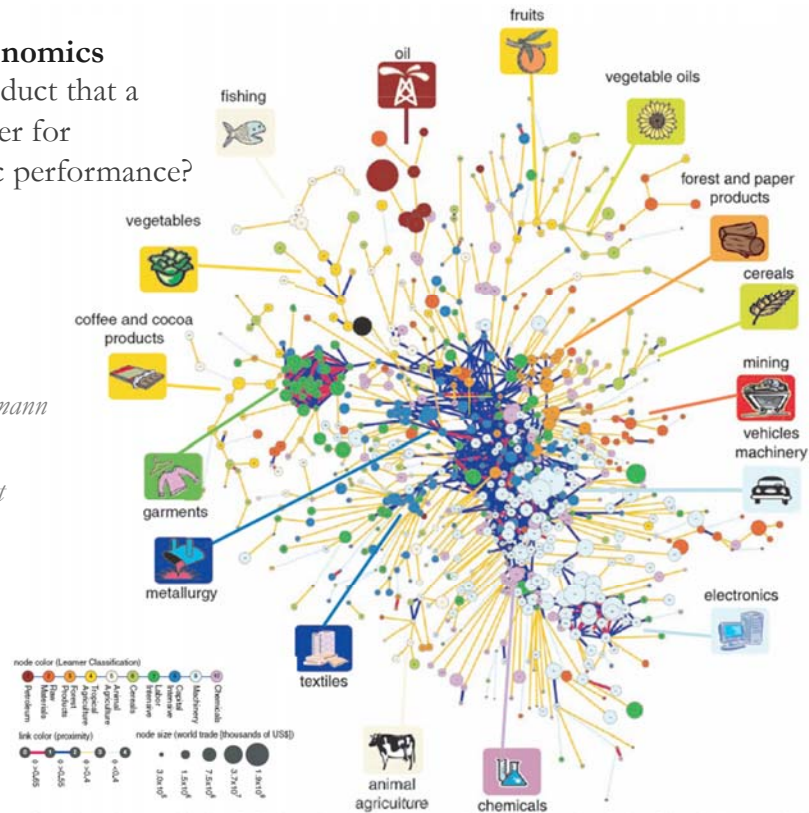
62



## Computational Economics

Does the type of product that a country exports matter for subsequent economic performance?

*C. A. Hidalgo, B. Klinger, A.-L. Barabási, R. Hausmann (2007) The Product Space Conditions the Development of Nations. Science 317, 482 (2007).*



**Fig. 1.** The product space. (A) Hierarchically clustered proximity matrix representing the 775 SITC-4 product classes exported in the 1998–2000 period. (B) Network representation of the product space. Links are color coded with their proximity value. The sizes of the nodes are proportional to world trade, and their colors are chosen according to the classification introduced by Leamer.

## Computational Social Science

Studying large scale social networks such as Wikipedia

*Second Sight: An Emergent Mosaic of Wikipedian Activity, The NewScientist, May 19, 2007*



### Second sight

Image: Bruce W. Heer and Todd M. Hollaway

#### Power struggle

How do you keep track of the bubbling mass of information that is Wikipedia? This chaotic-looking mosaic is one attempt to show which topics are contained in the online encyclopedia.



...pages at the time of writing include entries on Sheffield Wednesday football club, Mikhail Gorbachev and pigs). The mosaic has been commended in a competition for images that visualise network dynamics, coinciding with this week's International Workshop and Conference on Network Science in Bloomington.





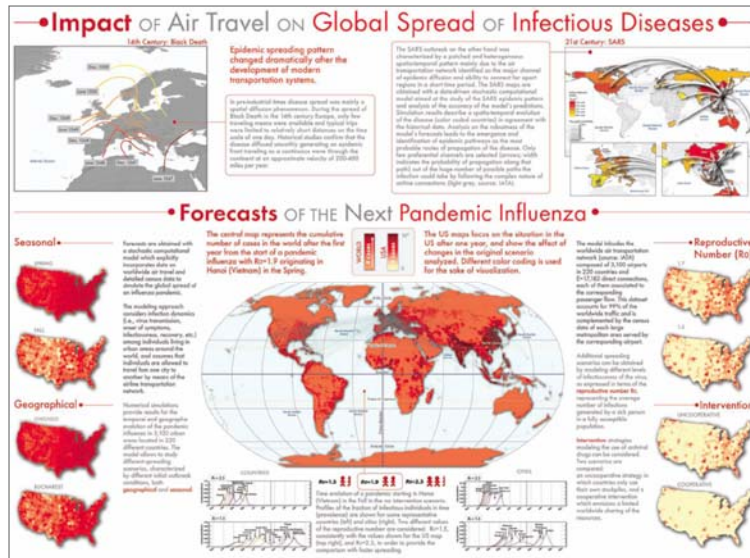
# Computational Epidemics

Forecasting (and preventing the effects of) the next pandemic.

*Epidemic Modeling in Complex realities*, V. Colizza, A. Barrat, M. Barthelemy, A. Vespignani, *Comptes Rendus Biologie*, 330, 364-374 (2007).

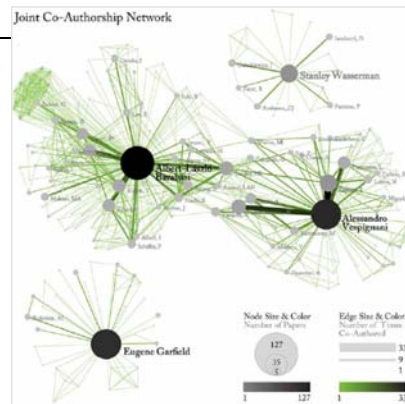
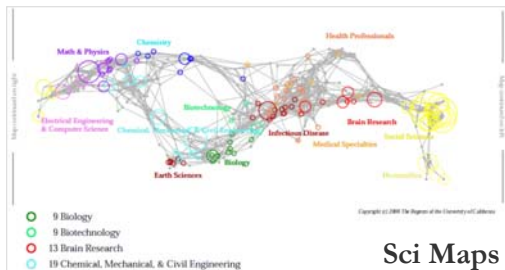
*Reaction-diffusion processes and metapopulation models in heterogeneous networks*, V. Colizza, R. Pastor-Satorras, A. Vespignani, *Nature Physics* 3, 276-282 (2007).

*Modeling the Worldwide Spread of Pandemic Influenza: Baseline Case and Containment Interventions*, V. Colizza, A. Barrat, M. Barthelemy, A.-J. Valleron, A. Vespignani, *PLoS-Medicine* 4, e13, 95-110 (2007).

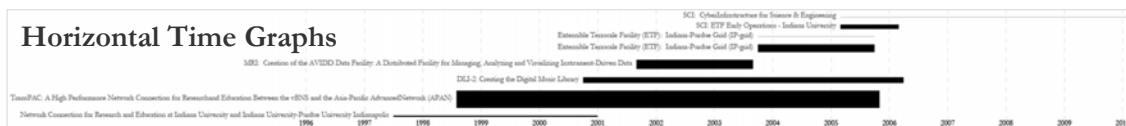


## Sci² Tool – “Open Code for S&T Assessment”

OSGi/CIShell powered tool with NWB plugins and many new scientometrics and visualizations plugins.



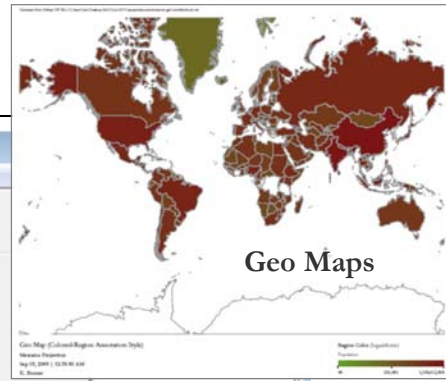
## Horizontal Time Graphs



Börner, Katy, Huang, Weixia (Bonnie), Linnemeier, Micah, Dubon, Russell Jackson, Phillips, Patrick, Ma, Nianli, Zoss, Angela, Guo, Hanning & Price, Mark. (2009). *Rete-Netzwerk-Red: Analyzing and Visualizing Scholarly Networks Using the Scholarly Database and the Network Workbench Tool*. *Proceedings of ISIS 2009: 12th International Conference on Scientometrics and Informetrics*, Rio de Janeiro, Brazil, July 14-17. Vol. 2, pp. 619-630.



## Sci² Tool Vis cont.



Sci² Tool

File Preprocessing Modeling Analysis Visualization Scientometrics Help

Console

Welcome to the Science of Science Tool (Sci²). The development of this tool is supported in Network Science center and the School of Li Indiana University, the National Science Foundation and IIS-0715303, and the James S. McDonnell Cyberinfrastructure portal (<http://sci.slis.indiana.edu>).

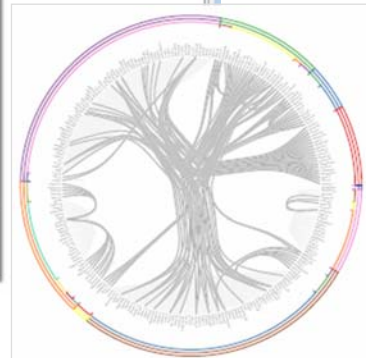
The primary investigators are Katy Börner, In SciTech Strategies Inc. The Sci² tool was developed by J. Duhon, Patrick A. Phillips, Chintan Tank, a Cyberinfrastructure Shell (<http://cishell.org>) for Network Science Center (<http://cns.slis.indiana.edu>). Many algorithm plugins were derived from the Network Science Center (<http://nwb.slis.indiana.edu>).

Please cite as follows:  
Sci² Team. (2009). Science of Science Tool. In SciTech Strategies Inc., <http://sci.slis.indiana.edu>.

Scheduler

Remove From List  Remove completed

!	Algorithm Name	Date	Time	% Con
<input checked="" type="checkbox"/>	Extract Co-Author Netw...	09/03/2009	00:15:20 AM	<div style="width: 100%;"></div>
<input checked="" type="checkbox"/>	Load and Clean ISI File	09/03/2009	00:15:05 AM	<div style="width: 100%;"></div>



67

Sci² Tool

A tool for science of science research & practice

Email Address

Password

Login

Forgot your password?

To recover your account password, please visit our [password recovery page](#).

Not registered yet?

[Register now](#)

Tutorials

Katy Börner (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.

- 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: [VIVO National Researcher Networking](#)
- Tutorial #12: [Future Developments](#)

<http://sci2.cns.in.edu>

<http://sci2.wiki.cns.in.edu>

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

68

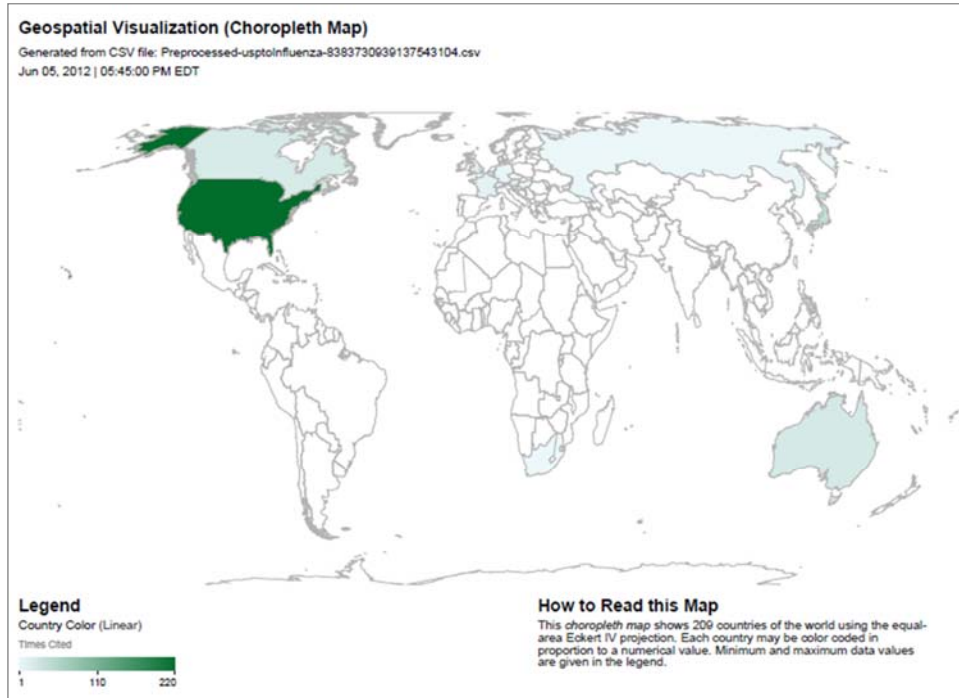




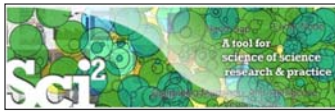




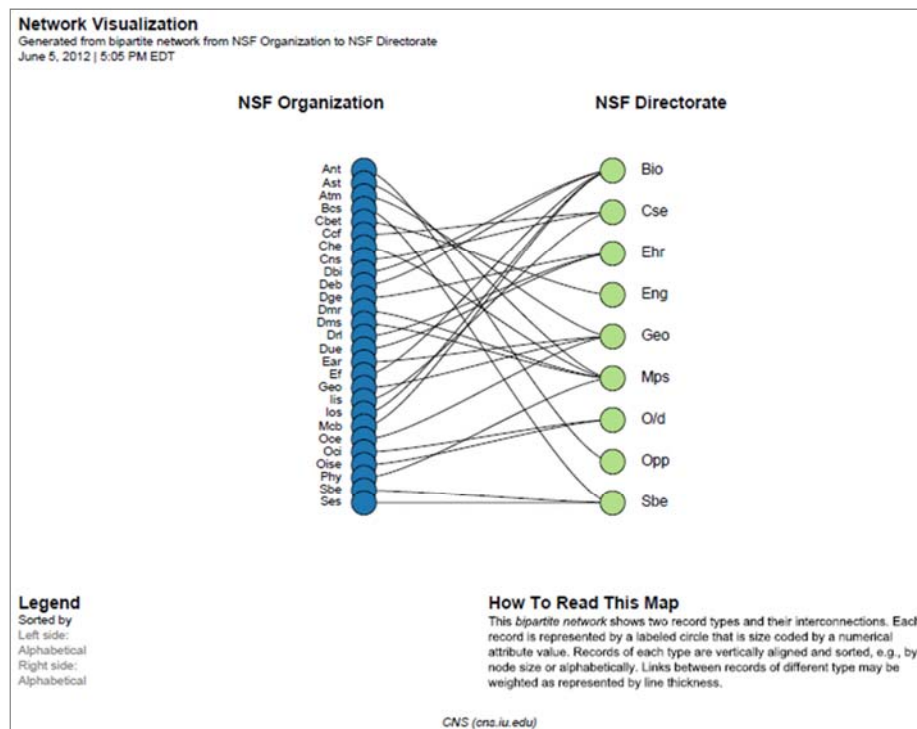
## Geospatial (Where): World Map Overlay



71



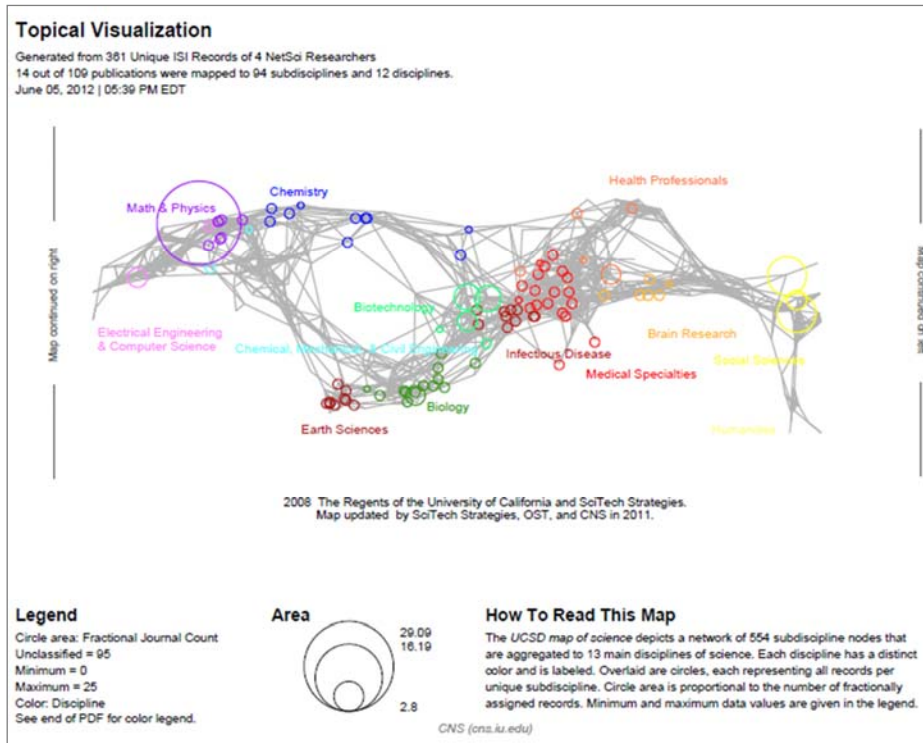
## Network (With Whom): Bipartite Network



72



# Topical (What): UCSD Map of Science Overlays



**EpiC Tool**  
 File | Compartmental Modeling | Networks | Simulation | Visualization | R | Help

Welcome to the EpiC tool, which supports the modeling, analysis, and visualization of epidemic processes.

The EpiC project (<http://epic.sls.indiana.edu>) is supported in part by the NIH RM-07-004 award. The primary investigators are Dr. Katy Börner, Dr. Alessandro Vespignani, and Dr. Jim Sherman.

Please cite as follows:  
 EpiC Team. (2009). EpiC Tool. Indiana University.

**Simulation** | Visualization | R | Help  
 Single-Population  
 Exact  
 Network

**Visualization** | R | Help  
 Line Graph

**R** | Help  
 Create an R Instance  
 Run Rgui  
 Import Table Into R  
 Export Table From R

**EpiC**  
 cyberinfrastructure for NETWORK SCIENCE CENTER  
 CIShell Powered

A number of other projects recently adopted OSGi and/or CIShell:

- USA**
- *Cytoscape* (<http://cytoscape.org>) Led by Trey Ideker at the University of California, San Diego is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating these interactions with gene expression profiles and other state data (Shannon et al., 2002).
  - *MAEviz* (<https://wiki.ncsa.uiuc.edu/display/MAE/Home>) Managed by Jong Lee at NCSA is an open-source, extensible software platform which supports seismic risk assessment based on the Mid-America Earthquake (MAE) Center research.
- Europe**
- *Taverna Workbench* (<http://taverna.org.uk>) Developed by the myGrid team (<http://mygrid.org.uk>) led by Carol Goble at the University of Manchester, U.K. is a free software tool for designing and executing workflows (Hull et al., 2006). Taverna allows users to integrate many different software tools, including over 30,000 web services.
  - *TEXTrend* (<http://texttrend.org>) Led by George Kampis at Eötvös Loránd University, Budapest, Hungary supports natural language processing (NLP), classification/mining, and graph algorithms for the analysis of business and governmental text corpuses with an inherently temporal component.
  - *DynaNets* (<http://www.dynanets.org>) Coordinated by Peter M.A. Sloot at the University of Amsterdam, The Netherlands develops algorithms to study evolving networks.
  - *SISOB* (<http://sisob.lcc.uma.es>) An Observatory for Science in Society Based in Social Models.
- As the functionality of OSGi-based software frameworks improves and the number and diversity of dataset and algorithm plugins increases, the capabilities of custom tools will expand.

75

## References

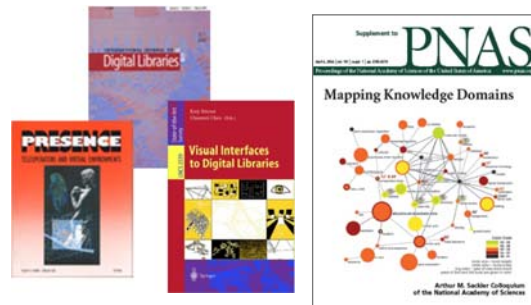
Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains**. In Blaise Cronin (Ed.), *ARIST*, Medford, NJ: Information Today, Volume 37, Chapter 5, pp. 179-255.  
<http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf>

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., Volume 41, Chapter 12, pp. 537-607.  
<http://ivl.slis.indiana.edu/km/pub/2007-borner-arist.pdf>

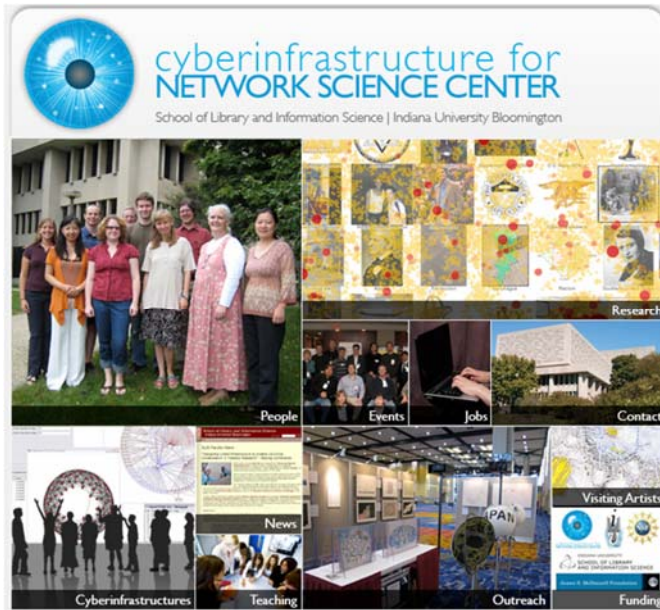
Börner, Katy (2010) **Atlas of Science**. MIT Press.  
<http://scimaps.org/atlas>

Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.



76





All papers, maps, tools, talks, press are linked from <http://cns.iu.edu>

CNS Facebook: <http://www.facebook.com/cnscenter>

Mapping Science Exhibit Facebook: <http://www.facebook.com/mappingscience>